Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 19 (2017), Issue 2   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Viewpoints

Original Paper

A Web-Based Physical Activity Intervention for Spanish-Speaking Latinas: A Costs and Cost-Effectiveness Analysis

Britta Larsen[1], PhD; Bess Marcus[1], PhD; Dori Pekmezi[2], PhD; Sheri Hartman[1], PhD; Todd Gilmer[1], PhD

[1]University of California, San Diego, Department of Family Medicine and Public Health, La Jolla, CA, United States

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

Corresponding Author:
Britta Larsen, PhD
University of California, San Diego
Department of Family Medicine and Public Health
9500 Gilman Drive
La Jolla, CA, 92093-0628
United States
Phone: 1 8585348363
Fax: 1 8585347517
Email: blarsen@ucsd.edu

Abstract

Background: Latinas report particularly low levels of physical activity and suffer from greater rates of lifestyle-related conditions such as obesity and diabetes. Interventions are needed that can increase physical activity in this growing population in a large-scale, cost-effective manner. Web-based interventions may have potential given the increase in Internet use among Latinas and the scalability of Web-based programs.

Objective: To examine the costs and cost-effectiveness of a Web-based, Spanish-language physical activity intervention for Latinas compared to a wellness contact control.

Methods: Healthy adult Latina women (N=205) were recruited from the community and randomly assigned to receive a Spanish-language, Web-based, individually tailored physical activity intervention (intervention group) or were given access to a website with content on wellness topics other than physical activity (control group). Physical activity was measured using the 7-Day Physical Activity Recall interview and ActiGraph accelerometers at baseline, 6 months (ie, postintervention), and 12 months (ie, maintenance phase). Costs were estimated from a payer perspective and included all features necessary to implement the intervention in a community setting, including staff time (ie, wages, benefits, and overhead), materials, hardware, website hosting, and routine website maintenance.

Results: At 6 months, the costs of running the intervention and control groups were US $17 and US $8 per person per month, respectively. These costs fell to US $12 and US $6 per person per month at 12 months, respectively. Linear interpolation showed that intervention participants increased their physical activity by 1362 total minutes at 6 months (523 minutes by accelerometer) compared to 715 minutes for control participants (186 minutes by accelerometer). At 6 months, each minute increase in physical activity for the intervention group cost US $0.08 (US $0.20 by accelerometer) compared to US $0.07 for control participants (US $0.26 by accelerometer). Incremental cost-per-minute increases associated with the intervention were US $0.08 at 6 months and US $0.04 at 12 months (US $0.16 and US $0.08 by accelerometer, respectively). Sensitivity analyses showed variations in staffing costs or intervention effectiveness yielded only modest changes in incremental costs.

Conclusions: While the Web-based physical activity intervention was more expensive than the wellness control, both were quite low cost compared to face-to-face or mail-delivered interventions. Cost-effectiveness ranged markedly based on physical activity measure and was similar between the two conditions. Overall, the Web-based intervention was effective and low cost, suggesting a promising channel for increasing physical activity on a large scale in this at-risk population.

ClinicalTrial: Clinicaltrials.gov NCT01834287; https://clinicaltrials.gov/ct2/show/NCT01834287 (Archived by WebCite at http://www.webcitation.org/6nyjX9Jrh)

(J Med Internet Res 2017;19(2):e43) doi:10.2196/jmir.6257

KEYWORDS

cost-effectiveness; physical activity; Latinos; Latinas; Web-based interventions; health disparities

Introduction

Physical activity is associated with a lower risk of heart disease, stroke, type 2 diabetes, depression, and some cancers [1]; however, only about 1 in 5 (21%) US adults meet the 2008 Federal Physical Activity Guidelines [2]. Recent analyses estimate that US $131 billion (95% CI $91 billion-$172 billion) of health care expenditures per year are associated with inadequate levels of physical activity. Moreover, mean annual health care expenditures are almost 30% higher among inactive adults compared to active adults [3]. There is a critical need for high-reach, cost-effective interventions to increase physical activity given the substantial public health burden and expenses related to such high levels of inactivity.

Internet-based interventions have great potential for widespread dissemination, thus many recent studies have focused on using this platform for physical activity promotion. In fact, a recent comprehensive review of Web-based physical activity interventions for adults identified 72 such programs, 44 of which (61%) reported significant increases in physical activity [4]. However, the review noted that this work was conducted in predominantly non-Latino, white samples and called for better representation of underserved populations, specifically racial and ethnic minorities, in future Internet-based physical activity intervention studies.

Latinos are the largest ethnic minority in the United States [5] and suffer marked health disparities. According to the US Census Bureau, Latinos (53 million) made up 17% of the US population growth between 2000 and 2010, accounting for more than half of the nation's population growth [6]. Latinos report particularly low rates of physical activity and are disproportionately affected by related health conditions, especially Latina women. Only 38.2% of Latinas in the United States report meeting the federal guidelines for performance of aerobic physical activity (ie, >150 minutes/week of at least moderate-intensity aerobic activity), which is markedly less than non-Latina white women (50.9 %) and Latino men (47.0 %). Moreover, Latinas are more likely to be obese than non-Latina white women and experience excess burden from inactivity-related conditions, such as diabetes and stroke [7].

Evidence suggests that Internet-delivered interventions may be cost-effective. Past research in Dutch adults over 50 years of age indicated that Web-delivered interventions are a cost-effective way to increase physical activity when compared to no intervention [8]. Internet-delivered interventions may represent a particularly appropriate low-cost approach to physical activity promotion in Latinas due to the Internet's ability to reach large numbers of people in the convenience of their own homes and address barriers to physical activity participation commonly cited by Latinas (eg, childcare and transportation) [9]. While in past years Internet use among Latinos was lower than among non-Latino whites, Internet use has grown markedly among Latinos in recent years, such that 81% of Latino adults report using the Internet compared to 85% of non-Latino white adults. Internet-based interventions, therefore, may be both low cost and have potential for broad dissemination in this population.

Our research team has developed a theory-based, individually tailored, Internet-delivered intervention [10,11] that was adapted for use in Latinas through extensive formative research—focus groups, cognitive interviews, and pilot studies—and tested in a randomized controlled trial [12]. Increases in minutes per week of at least moderate-intensity aerobic physical activity were significantly greater in the Internet intervention arm compared to the wellness contact control arm at 6 months [13] and were largely sustained at 12 months. However, to date there have been no cost or cost-effectiveness analyses conducted of such Internet-based physical activity interventions in Latinas. Therefore, the purpose of this paper is to assess the costs of this linguistically and culturally adapted Internet-based physical activity intervention for Latinas compared to the costs of a Web-based contact-matched control group. The purpose is also to evaluate the cost-effectiveness of this intervention for increasing moderate- to vigorous-intensity physical activity (MVPA) in Latinas.

Methods

Design

The Pasos Hacia La Salud study was a randomized controlled trial of an individually tailored, Spanish-language, Web-based physical activity intervention compared to a Web-based wellness control. Participants in the intervention condition group were given access to a website with individually tailored physical activity information; they also received monthly, personalized reports for 12 months. Participants in the control condition group were given access to a website with information on wellness topics other than physical activity (eg, diet and stress reduction); they also received alerts to access new materials with the same frequency as the intervention group. The primary outcome was weekly minutes of MVPA at 6 months measured by the 7-Day Physical Activity Recall (PAR) interview. Activity was also measured by accelerometer. A secondary outcome was minutes of MVPA at 12 months.

Setting and Sample

Participants included 205 adult Latinas recruited from the community in San Diego County, California. Eligible participants were between 18 and 65 years of age, underactive (ie, engaging in less than 60 minutes per week of MVPA), and self-identified as Latina and/or Hispanic women. The study focused on women because our formative research showed that common barriers, motivators, and activity preferences were markedly different between Latino men and Latina women. In addition, physical inactivity is much more common among Latina women than among Latino men [14]. Exclusion criteria included current or planned pregnancy, plans to move from the area within 12 months, a body mass index of 45 kg/m^2 or greater, and any health condition that might make unsupervised physical activity unsafe as determined by the Physical Activity Readiness Questionnaire [15], including history of heart disease, stroke, diabetes, or orthopedic problems. Participants also had to be willing to be randomized to either of the two conditions. A detailed breakdown of participant eligibility and flow diagram of participant allocation has been published elsewhere [12].

Human subjects research approval was granted by the University of California, San Diego, Human Research Protections Program, and all participants gave written informed consent.

Intervention

The intervention was based on the Transtheoretical Model and Social Cognitive Theory. Participants filled out monthly online surveys about physical activity, cognitive and behavioral strategies to change behavior, self-efficacy, and other psychosocial constructs. Responses to these surveys were used to generate individually tailored reports for each participant, with feedback on changes over time and information about how their answers compared to those of other active women. Participants also received online physical activity manuals, which were matched to their readiness for changing physical activity behavior. Other features of the website included a calendar for goal setting and logging daily minutes of activity and steps, a message board for interacting with other participants, an *ask the expert* page, and a guide to local free and low-cost physical activity resources. Participants received regular emails with tip sheets on topics such as finding time to exercise, staying motivated, and other topics highlighted in formative research as being important to this population, such as childcare and cultural norms. An initial on-site visit was held to conduct a goal-setting session and orient participants to the website. At this initial visit, participants also received a pedometer for entering their daily steps on the website activity-tracking calendar, a binder with physical activity information sheets, a music CD, staff contact information, and the option of taking home a Spanish-language exercise DVD.

The first 6 months comprised the intensive intervention stage. Each month participants filled out questionnaires and received a tailored report and stage-matched online manual. Participants also received regular emails prompting them to access new information sheets and other materials on the website—weekly emails in month 1, biweekly emails in months 2 and 3, and monthly emails in months 4-6. All participants received a phone call after 1 week to help with pedometer and website use; received another call at 1 month to check in, answer questions, and help with goal setting; and participated in a repeat goal-setting session at 6 months. The second 6 months comprised a maintenance phase, during which participants received a monthly email prompt to visit the website to fill out the questionnaire to generate a personalized report. A final call to check in was made at month 9. The website and all materials were in Spanish.

Wellness Contact Control

The control group received access to a Spanish-language website with a similar look and feel as the physical activity-based intervention site. However, this site included information on health topics other than physical activity, including diet, stress reduction, and sleep. In order to control for contact time, participants also engaged in an initial visit to be oriented to their website, received emails on the same schedule as the intervention group with new information sheets on various wellness topics, and filled out monthly questionnaires on wellness topics. Like those in the intervention condition group, they also received a phone call to check in at months 1 and 9, and had a short site visit at month 6.

Measures

Clinical Outcome Measures

The primary outcome, upon which the study was powered, was change in MVPA from baseline to 6 months as measured by the 7-Day PAR. The 7-Day PAR is a self-report measure administered by trained, certified interviewers that asks participants to report the amount of weekly minutes spent in activities of different intensities—light, moderate, hard, very hard, and sleep—across a range of settings and activity types (eg, leisure, transport, and occupational). This measure has shown acceptable reliability and congruent validity with more objective physical activity measures and shows sensitivity to changes over time [16,17]. Staff members performing the 7-Day PAR were blinded to condition.

As an additional primary outcome, participants also wore ActiGraph GT3X+ accelerometers for the week corresponding to the self-report measure at the same three time points—baseline, 6 months, and 12 months. A minimum count of 1952 was set as a threshold for MVPA and a minimum bout duration of 10 minutes was used. Valid wear time was considered as at least 10 hours of wear on at least 5 days or at least 3000 minutes of wear time on at least 4 days.

Costs

Costs were estimated from a payer perspective and included all costs necessary to deliver the developed Web-based intervention in a clinical or community setting. This included staff time for training and delivering the intervention (ie, salary, benefits, and overhead) and cost of website maintenance and materials based on actual costs incurred during the trial. Costs associated exclusively with research activities, such as baseline and follow-up assessments, obtaining consent, and participant compensation, were not included. We also did not include costs of developing the website, as delivering the intervention would utilize the existing website and would not require further development. Costs associated with maintaining the website, such as Web hosting and technical support, were included.

Personnel Time

Costs for personnel were calculated by multiplying the time required for specific tasks—for the intervention and control groups separately—by standard University of California salary and fringe rates. Personnel time was determined by asking research staff to log the amount of time spent on nonresearch-related activities. This included training time for both the trainer and trainee; conducting initial baseline goal-setting visits and 6-month visits; scheduling and conducting 1-week, 1-month, and 9-month calls, including time for failed contact attempts; and time to compile study materials, such as the binder and CD.

Staff time also included routine maintenance of the websites, including checking the message board for appropriate comments and responding to *ask the expert* questions. It also included time needed to resend messages and materials, such as replacing lost pedometers or sending emails to corrected email addresses.

Staff costs were based on standard published salaries at the University of California, San Diego, for research staff qualified for each task. A masters- or PhD-level trainer's annual salary was US $60,000 plus 44% benefits or US $86,520 annually (US $43.26 per hour) and a research associate with a bachelor's degree received US $38,941 annual salary plus 44% benefits or US $56,153 annually (US $28.08 per hour). These hourly costs were increased by 10% to account for overhead costs of shared space.

Website Costs

Costs associated with using the website to deliver the intervention in a community or clinic setting included Web hosting and regular IT support. Web hosting was priced at US $75 per month for each website. Technical support was estimated at 2 hours per month at a standard rate of US $95 per hour. These costs were based on actual costs charged by the Web developer (Illumina Interactive) during the trial once the website was developed and being used by participants. Costs were equal for the intervention and control group websites.

Materials

Material costs were based directly on what was actually spent in the trial. Materials included the study binders, music CDs, labels, business cards, and paper and ink for tip sheets in the initial study binder. Costs for these came from standard wholesale office supply prices. Those in the intervention arm also received an Accusplit pedometer at a cost of US $12.50 each. A video library of approximately 15 DVDs was also available for participants to borrow from; the DVDs cost approximately US $10 each.

Hardware

Hardware costs included a standard desktop computer and a printer for printing tip sheets for the study binders. Hardware costs were estimated using market prices in June 2014; costs were depreciated using a straight-line method assuming a 5-year depreciation period and 3 years of use in the study. Hardware costs were equal for the intervention and control groups.

Analysis

Costs were calculated as the total of all materials, hardware, personnel time including overhead and fringe benefits, website hosting, and website maintenance needed to run the study with the given number of participants. Research and development costs were not included. Change in total physical activity over the course of the study was calculated using linear interpolation of minutes from baseline to 6 months, then 6 months to 12 months, subtracting baseline physical activity. This was done for both subjective (7-Day PAR) and objective (accelerometer) measures of MVPA. Increase in physical activity was calculated using unadjusted mean weekly minutes at each time point for participants completing the trial (172/205, 83.9%). Dropout was equal across arms and there were no significant differences in any measures at baseline between those who did and did not complete the trial. There was also no significant difference between adjusted and unadjusted mean differences in MVPA. Cost-effectiveness was defined as the cost-per-minute increase in activity in each arm. This was determined by calculating the total cost per person at each time point, over the first 6 months and over the whole 12 months, and dividing this by the average total increase in physical activity across each time period. This was done separately for each condition in order to compare cost-effectiveness between the intervention and control groups. Finally, the incremental cost was defined as the additional cost-per-minute increase in the intervention condition beyond the change in MVPA in the control group; this was calculated by dividing the difference in change in MVPA between the two conditions by the difference in cost between the two conditions.

Sensitivity analyses were conducted by determining how cost-effectiveness, specifically incremental costs, would be impacted if personnel costs or intervention effectiveness increased or decreased by 20%. Because the cost per person could also vary depending on the number of participants enrolled, we also modeled how costs would change with varying participant numbers, multiplying the cost per person by the number of participants, but keeping fixed costs the same. Fixed costs included hardware (eg, computer and printer), technical support and Web hosting, and training costs. Variable costs included materials (eg, binders, paper, and pedometers) and the time for staff to deliver the intervention.

Results

Costs

Costs associated with running the intervention and control groups are presented in Table 1. Total cost of conducting the intervention at 6 months was US $10,712, which corresponded to a cost of US $17.17 per person per month. In contrast, the cost of running the control condition at 6 months was US $4900, or US $8.09 per person per month.

Table 1. Costs of study components for the intervention and control groups for the first 6 months (intensive intervention) and across the whole 12-month study period (N=205).

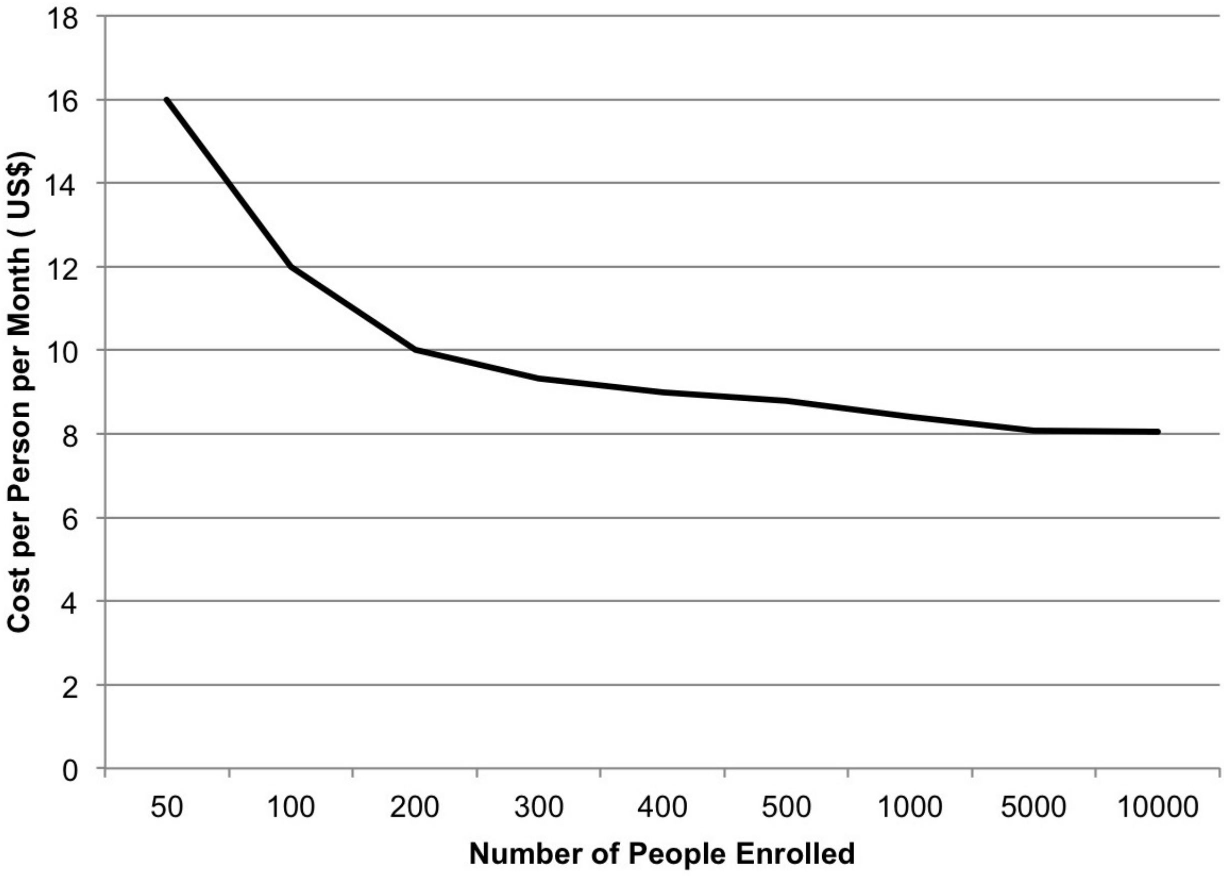
Item		Costs for intervention group (n=104), US $		Costs for wellness control group (n=101), US $	
		6 months	12 months (cumulative)	6 months	12 months (cumulative)
Personnel					
	Training	$471	$471	$157	$157
	Intervention delivery	$5674	$8083	$1976	$3119
Website					
	Maintenance	$270	$490	$8	$10
	Hosting	$450	$900	$450	$900
	Technical support	$1140	$2280	$1140	$2280
Hardware					
	Computer	$420	$420	$420	$420
	Printer	$240	$240	$240	$240
Materials					
	Pedometers	$1300	$1300	N/A[a]	N/A
	Paper, ink, binders, etc	$597	$597	$510	$510
	Videos	$150	$150	N/A	N/A
Total cost		$10,712	$14,781	$4900	$7634
Average cost per participant		$103	$142	$49	$76
Average cost per participant per month		$17	$12	$8	$6

[a] N/A: not applicable.

Cumulative total costs at 12 months were US $14,781 for the intervention group and US $7634 for the control group. Compared to costs at 6 months, this corresponded to a lower monthly cost of US $11.84 per participant per month for the intervention group and US $6.30 per participant per month for the control group. The largest expense for both conditions was personnel time, which was primarily devoted to conducting initial baseline visits, followed by IT support. The cost of providing pedometers in the intervention condition was also relatively high.

Figure 1 shows the cost to deliver the intervention per person over the entire 12 months modeled by number of people enrolled. With only 50 people enrolled, the cost per person per month was US $16; this dropped to US $12 with 100 people and US $10 with 200 people. Costs dropped off more slowly with 300 people or more, reaching an asymptote near US $8 per person per month.

Figure 1. Cost per person per month over 12 months by number of participants enrolled.



Clinical Outcome Measures

As reported elsewhere [13], MVPA increased from baseline to 6 months in both conditions. This change was significantly greater in intervention group participants, who increased from an average of 8.0 (SD 15.0) minutes of self-reported MVPA per week at baseline to an average of 112.8 (SD 97.1) minutes at 6 months. Baseline minutes were similar in control group participants, which were an average of 8.5 (SD 14.6) minutes per week, but only increased to an average of 63.5 (SD 88.7) minutes per week at 6 months. Assuming a linear increase in MVPA across the first 6 months of the study, this corresponds to a total increase of 1362 minutes of MVPA per person in the intervention group versus 715 minutes of MVPA per person for the control group.

Changes recorded by objective measures were smaller, but still significantly greater in the intervention group than the control group. Intervention group participants increased from an average of 35.8 (SD 69.7) minutes per week at baseline to an average of 75.8 (SD 91.0) minutes per week at 6 months. Control group participants increased from an average of 28.7 (SD 48.2) minutes per week at baseline to an average of 43.0 (SD 60.9) minutes per week at 6 months. Using linear interpolation assuming a linear increase in minutes, this corresponds to a total increase of 523.0 minutes of MVPA per person in the intervention group versus 186.3 minutes of MVPA per person for the control group.

Increases in MVPA were largely sustained at 12 months. On average, intervention group participants reported an average of 108.6 (SD 107.2) weekly minutes of self-reported MVPA at 12 months compared to an average of 75.9 (SD 89.8) minutes in the control group. Assuming a linear change from baseline to 6 months and from 6 months to 12 months, this corresponds to a total increase of 4032.6 minutes per person in the intervention group over the course of the study and 2306.2 minutes per person in the control group. Weekly minutes recorded by accelerometers at 12 months were an average of 70.4 (SD 86.4) for the intervention group participants versus an average of 55.5 (SD 74.6) for control group participants. This corresponds to an increase of 1496 minutes per person over the course of the study for the intervention group and 695 minutes per person for the control group.

Cost-Effectiveness

Table 2 Table 2 provides the estimates of cost-per-minute increases in MVPA in the intervention and control arms. Over 6 months, costs for the intervention arm were US $0.08 per minute increase per person for self-reported minutes and US $0.20 per minute per person for accelerometer-recorded minutes of MVPA. For the control group, these costs were US $0.07 and US $0.26 per minute increase per person, respectively. Costs per minute decreased for both arms over the whole 12 months: after 1 year, the intervention group costs were US $0.04 per minute increase per person for self-reported minutes and US $0.10 per minute increase per person for accelerometer-measured minutes of MVPA. For the control group, these costs were US $0.03 and US $0.11 per minute increase per person, respectively.

Table 2. Cost-per-minute increases in physical activity during the initial 6-month intervention and the full 12-month study (cumulative).

		Intervention group		Wellness control group	
		7-Day Physical Activity Recall	ActiGraph accelerometer	7-Day Physical Activity Recall	ActiGraph accelerometer
Total MVPA[a] minutes per person, minutes					
	Baseline to 6 months	1362	523	715	186
	Total 12 months	4033	1496	2306	696
MVPA minutes per person per month, minutes					
	Baseline to 6 months	227	87	119	31
	Total 12 months	336	125	192	58
Cost-per-minute increase in MVPA, US $					
	Baseline to 6 months	$0.08	$0.20	$0.07	$0.26
	Total 12 months	$0.04	$0.10	$0.03	$0.11
Incremental cost-per-minute increase in MVPA, US $					
	Baseline to 6 months	$0.08	$0.16	N/A[b]	N/A
	Total 12 months	$0.04	$0.08	N/A	N/A

[a] MVPA: moderate- to vigorous-intensity physical activity.

[b] N/A: not applicable.

Incremental costs for increases in MVPA (ie, cost-per-minute increases in the intervention group beyond that seen in the control group) varied by time point and measurement method. At 6 months, incremental costs were US $0.08 per minute for the 7-Day PAR and US $0.16 per minute for accelerometer-measured MVPA. At 12 months, these were US $0.04 per minute and US $0.08 per minute, respectively.

Sensitivity analyses (see Table 3) examined how changes in staffing costs and intervention effectiveness would influence cost-effectiveness, specifically incremental costs. A 20% increase in staffing costs yielded an increased incremental cost at 6 months from US $0.08 to US $0.10 per minute—US $0.16 to US $0.19 for accelerometers—while a 20% decrease in staffing costs yielded a decrease to US $0.07 and US $0.014 per minute for subjective and accelerometer-measured minutes, respectively. Variations in intervention effectiveness (ie, changes in MVPA minutes) yielded nearly identical changes. For all sensitivity analyses, changes in incremental costs were quite small at 12 months.

Table 3. Sensitivity analyses for incremental cost-per-minute increases in moderate- to vigorous-intensity physical activity in the intervention versus control groups.

Cost and effectiveness measures		6-month incremental cost per minute, US $		12-month incremental cost per minute, US $	
		7-Day Physical Activity Recall	ActiGraph accelerometer	7-Day Physical Activity Recall	ActiGraph accelerometer
Standard calculation		$0.08	$0.16	$0.04	$0.08
Staffing costs					
	+20%	$0.10	$0.19	$0.04	$0.10
	-20%	$0.07	$0.14	$0.03	$0.07
Intervention effectiveness					
	+20%	$0.07	$0.13	$0.03	$0.07
	-20%	$0.11	$0.20	$0.05	$0.12

Discussion

Principal Findings

Our Internet-based physical activity intervention for Spanish-speaking Latinas, which was previously shown to be effective for increasing MVPA, was also low cost with a cost of just US $12 per person per month over the whole 12-month period. One of the key benefits of an Internet-based intervention is the ability to scale the intervention for larger samples. Results showed that the cost per participant in the intervention group could be reduced from US $12 down to US $8 per month, with only about 300 users in the intervention arm. The most expensive cost by far was personnel time. Interventions that utilize automated systems that require little to no staff time are important for keeping the cost of physical activity programs low. This study utilized an automated system while still personalizing intervention materials. The greater cost of running the intervention condition was primarily driven by the increased staff time required to orient participants to the website, perform an individual goal-setting tutorial, and make check-in calls.

Although the monthly cost of the intervention arm was more expensive than the control arm, participants in the intervention arm increased their physical activity significantly more than those in the control arm. Consequently, the cost per minute increase was similar between the two groups, with the control arm being slightly more cost-effective for subjectively measured minutes and the intervention arm being slightly more cost-effective for objectively measured minutes. These results suggest that this Web-based intervention could produce significantly greater increases in MVPA at a similar price per minute. Given the vast benefits of physical activity and the low incremental cost of just US $0.04 per minute beyond that seen in the control arm, this Web-based intervention is a cost-effective approach to health promotion.

Comparison With Prior Work

TThe costs of this Web-based intervention are considerably lower than a printed mail-delivered version of this program completed with Latina women, which cost US $17 per person per month [18]. The mail-delivered intervention was slightly less effective and more costly, thus the cost per minute at 6 months was also significantly higher in the mail-delivered program (US $0.18) than in the Web-based program (US $0.08). These costs would likely diverge farther with wider dissemination, as many of the costs of the Internet intervention do not increase with increased users, while the print version requires postage, printing materials, mailing materials, and staff time to process tailored questionnaires for each person.

It is difficult to compare these costs to other Web-based interventions, as few Internet-based physical activity interventions have examined cost-effectiveness and none could be found that examined it with underserved populations. Additionally, across cost-effectiveness studies there is a lack of homogeneity regarding which variables are included when calculating costs, preventing direct comparisons between studies. For example, one paper comparing two Web-based studies reported much lower average costs for conducting their interventions than this study; however, no staff time was included [8]. In addition, while some analyses included recruitment costs, we did not include them in this study as recruitment costs incurred during the trial were for a research study and were therefore not representative of recruitment costs for delivering the intervention in a real-world setting. Some recruitment costs would likely be necessary in a clinical or community setting, though these would vary greatly depending on whether they targeted an existing patient or membership base or the community at large.

On that note, costs reported here should be seen as a guide to potential delivery costs rather than an exact estimate, as costs would likely vary widely depending on delivery site. The intervention examined here could potentially be implemented in a health system or community center, each of which would have vastly different costs for staffing, recruitment (if any), and overhead. Sensitivity analyses showed that cost-effectiveness, measured by incremental costs, would remain relatively stable despite ranges in staffing costs. Larger variations were seen in cost-effectiveness when examining change in intervention effectiveness, which would likely range greatly across populations. Variations in cost-effectiveness were perhaps modeled best in this study by comparing subjective and objective measures of physical activity, which showed marked differences in activity gains. These differences in reported minutes could be due to subjective reporting bias or inaccurate recall of duration or intensity of activities. Some discrepancy between them should be expected, however, as accelerometry and the 7-Day PAR measure different behaviors (ie, activity vs hip acceleration). Accelerometry is also unable to accurately measure some activities, such as biking and swimming, and underestimates energy expenditure for some activities, such as hiking and walking on an incline or activities using upper body movement [19].

As the purpose of this study was to estimate the cost of delivering the already-developed intervention, we did not take into account the cost associated with developing the intervention. A previous study with an Internet-delivered intervention for adults did examine the development costs [20] and determined that although developing the intervention cost about US $100,000, only 352 participants would be needed to overcome the start-up costs of the Internet intervention relative to delivering the print-based version. Since start-up costs only occur once and may be easily mitigated, our focus on cost of dissemination is important as some costs, such as pedometers, do not decrease with an increased number of participants.

Strengths and Limitations

One limitation is that we did not consider costs for updating the website over time. As technology grows and changes it is important to have a website that changes over time as well, which may incur costs over that of regular maintenance and technical support. Over the 3 years of running this intervention, the main modifications to the website were updating resources, which study staff were able to do; we were not able to estimate changes that would be of significant cost to updating the website in future years. The Web-based contact control group allowed for an assessment of intervention effectiveness, but did not allow for a comparison of costs across delivery modalities. Additionally, this analysis estimated costs from the payer perspective, but was unable to estimate costs from a societal perspective. We also did not include recruitment costs, as those associated with the trial would not be a good estimate of recruitment costs in a nonresearch setting.

This study has a number of strengths, including using data from a randomized controlled trial and relying on current market data for materials, equipment, and personnel costs. This study also targeted Spanish-speaking Latinas and demonstrated cost-effectiveness at the end of the intervention (ie, 6 months) and during the maintenance phase (ie, 12 months). Additionally, cost-effectiveness was calculated using both self-reported and objectively measured physical activity.

Given the disparities in chronic diseases among Latinas, next steps should be to determine whether physical activity interventions could reduce health care costs. This study focused on relatively short-term physical activity outcomes and did not collect information on health care costs. Future research that includes health care costs could provide a more comprehensive understanding of the potential costs and quality-adjusted life years that could be gained from increasing physical activity in Latinas [21]. As health habits in Latino families are more closely related than in non-Latino white families, particularly among mothers and children [22], future research should also investigate how increasing physical activity in Latina women could influence health and health care costs for their families and communities.

Conclusions

Results from this study indicated that a tailored, Internet-delivered intervention is a cost-effective approach to increasing physical activity among Latinas that has the potential for dissemination. Using the Internet allowed for the delivery cost of the intervention to be cut in half as the number of participants using the intervention increased. Large-scale implementations of interventions that demonstrate cost-effectiveness have the potential to reduce health disparities, benefiting public health.

Acknowledgments

This work was supported by the National Cancer Institute, National Institutes of Health (grant number: 5R01CA159954). We would like to thank the staff that made the study possible, including Raul Fortunet, Karla Nuñez, Rachelle Edgar, Madison Noble, Daniah Tanori, David Bakal, and Dr Veronica Villarreal at the University of California, San Diego, and Michael Getz at Illumina Interactive.

Conflicts of Interest

None declared.

References

1. health.gov. Rockville, MD: Office of Disease Prevention and Health Promotion; 2016. Physical Activity Guidelines URL: https://health.gov/paguidelines/ [accessed 2017-02-10] [WebCite Cache ID 6iSjNi5Wg]
2. Centers for Disease Control and Prevention (CDC). Adult participation in aerobic and muscle-strengthening physical activities--United States, 2011. MMWR Morb Mortal Wkly Rep 2013 May 03;62(17):326-330 https://www.cdc.gov/mmwr/pdf/wk/mm6217.pdf [FREE Full text] [Medline: 23636025]
3. Carlson SA, Fulton JE, Pratt M, Yang Z, Adams EK. Inadequate physical activity and health care expenditures in the United States. Prog Cardiovasc Dis 2015;57(4):315-323 [FREE Full text] [doi: 10.1016/j.pcad.2014.08.002] [Medline: 25559060]
4. Joseph RP, Durant NH, Benitez TJ, Pekmezi DW. Internet-based physical activity interventions. Am J Lifestyle Med 2014 Dec;8(1):42-68 [FREE Full text] [doi: 10.1177/1559827613498059] [Medline: 25045343]
5. Humes KR, Jones NA, Ramirez RR. Overview of Race and Hispanic Origin: 2010. Suitland, MD: US Census Bureau; 2011 Mar. URL: http://www.census.gov/content/dam/Census/library/publications/2011/dec/c2010br-02.pdf [accessed 2017-02-10] [WebCite Cache ID 6oAwR8HT5]
6. Passel JS, Cohn D, Lopez MH. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2011 Mar 24. Hispanics account for more than half of nation's growth in past decade URL: http://www.pewhispanic.org/2011/03/24/hispanics-account-for-more-than-half-of-nations-growth-in-past-decade/ [accessed 2016-06-22] [WebCite Cache ID 6iSkrJc2L]
7. Schiller JS, Lucas JW, Ward BW, Peregoy JA. Summary health statistics for US adults: National Health Interview Survey, 2010. Vital Health Stat 10 2012 Jan(252):1-207 [FREE Full text] [Medline: 22834228]
8. Golsteijn RH, Peels DA, Evers SM, Bolman C, Mudde AN, de Vries H, et al. Cost-effectiveness and cost-utility of a Web-based or print-delivered tailored intervention to promote physical activity among adults aged over fifty: An economic evaluation of the Active Plus intervention. Int J Behav Nutr Phys Act 2014;11:122 [FREE Full text] [doi: 10.1186/s12966-014-0122-z] [Medline: 25262435]
9. Martinez SM, Arredondo EM, Perez G, Baquero B. Individual, social, and environmental barriers to and facilitators of physical activity among Latinas living in San Diego County: Focus group results. Fam Community Health 2009;32(1):22-33. [doi: 10.1097/01.FCH.0000342814.42025.6d] [Medline: 19092432]
10. Marcus BH, Lewis BA, Williams DM, Dunsiger S, Jakicic JM, Whiteley JA, et al. A comparison of Internet and print-based physical activity interventions. Arch Intern Med 2007 May 14;167(9):944-949. [doi: 10.1001/archinte.167.9.944] [Medline: 17502536]
11. Marcus BH, Lewis BA, Williams DM, Whiteley JA, Albrecht AE, Jakicic JM, et al. Step into Motion: A randomized trial examining the relative efficacy of Internet vs print-based physical activity interventions. Contemp Clin Trials 2007 Nov;28(6):737-747. [doi: 10.1016/j.cct.2007.04.003] [Medline: 17616486]
12. Marcus BH, Hartman SJ, Pekmezi D, Dunsiger SI, Linke S, Marquez B, et al. Using interactive Internet technology to promote physical activity in Latinas: Rationale, design, and baseline findings of Pasos Hacia La Salud. Contemp Clin Trials 2015 Aug 5:149-158. [doi: 10.1016/j.cct.2015.08.004] [Medline: 26255237]

13. Marcus BH, Hartman SJ, Larsen BA, Pekmezi D, Dunsiger SI, Linke S, et al. Pasos Hacia La Salud: A randomized controlled trial of an Internet-delivered physical activity intervention for Latinas. Int J Behav Nutr Phys Act 2016;13(1):62 [FREE Full text] [doi: 10.1186/s12966-016-0385-7] [Medline: 27234302]
14. Blackwell DL, Lucas JW, Clarke TC. Summary health statistics for US adults: National health interview survey, 2012. Vital Health Stat 10 2014 Feb(260):1-161 [FREE Full text] [Medline: 24819891]
15. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
16. Hayden-Wade HA, Coleman KJ, Sallis JF, Armstrong C. Validation of the telephone and in-person interview versions of the 7-day PAR. Med Sci Sports Exerc 2003 May;35(5):801-809. [doi: 10.1249/01.MSS.0000064941.43869.4E] [Medline: 12750590]
17. Leenders NY, Sherman WM, Nagaraja HN, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. Med Sci Sports Exerc 2001 Jul;33(7):1233-1240. [Medline: 11445774]
18. Larsen B, Gilmer T, Pekmezi D, Napolitano MA, Marcus BH. Cost effectiveness of a mail-delivered individually tailored physical activity intervention for Latinas vs a mailed contact control. Int J Behav Nutr Phys Act 2015;12:140 [FREE Full text] [doi: 10.1186/s12966-015-0302-5] [Medline: 26559336]
19. Hendelman D, Miller K, Baggett C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S442-S449. [Medline: 10993413]
20. Lewis BA, Williams DM, Neighbors CJ, Jakicic JM, Marcus BH. Cost analysis of Internet vs print interventions for physical activity promotion. Psychol Sport Exerc 2010 May 1;11(3):246-249 [FREE Full text] [doi: 10.1016/j.psychsport.2009.10.002] [Medline: 20401164]
21. Peels DA, Hoogenveen RR, Feenstra TL, Golsteijn RH, Bolman C, Mudde AN, et al. Long-term health outcomes and cost-effectiveness of a computer-tailored physical activity intervention among people aged over fifty: Modelling the results of a randomized controlled trial. BMC Public Health 2014 Oct 23;14:1099 [FREE Full text] [doi: 10.1186/1471-2458-14-1099] [Medline: 25342517]
22. Sallis JF, Patterson TL, Buono MJ, Atkins CJ, Nader PR. Aggregation of physical activity habits in Mexican-American and Anglo families. J Behav Med 1988 Feb;11(1):31-41. [Medline: 3367370]

Abbreviations

MVPA: moderate- to vigorous-intensity physical activity
N/A: not applicable
PAR: Physical Activity Recall

Edited by G Eysenbach; submitted 22.06.16; peer-reviewed by N Crespo, S Buchholz; comments to author 17.11.16; revised version received 21.12.16; accepted 14.01.17; published 22.02.17.

Please cite as:
Larsen B, Marcus B, Pekmezi D, Hartman S, Gilmer T
A Web-Based Physical Activity Intervention for Spanish-Speaking Latinas: A Costs and Cost-Effectiveness Analysis
J Med Internet Res 2017;19(2):e43
URL: http://www.jmir.org/2017/2/e43/
doi:10.2196/jmir.6257
PMID:28228368

Original Paper

Commencing and Persisting With a Web-Based Cognitive Behavioral Intervention for Insomnia: A Qualitative Study of Treatment Completers

Charles Chan[1], MBBS, MPsych, FRANZCP; Stacey West[1], BPsych, PGDip Psych, MBMSc; Nick Glozier[1], MBBS, FRANZCP, PhD

Brain and Mind Centre, University of Sydney, Camperdown, Australia

Corresponding Author:
Charles Chan, MBBS, MPsych, FRANZCP
Brain and Mind Centre
University of Sydney
Professor Marie Bashir Centre Level 5
Building 11 RPAH 67-73 Missenden Road
Camperdown, 2050
Australia
Phone: 61 02 9966 7408
Fax: 61 02 9789 3344
Email: sccchan@gmail.com

Abstract

Background: Computerized cognitive behavioral therapy for insomnia (CCBT-I) has a growing evidence base as a stand-alone intervention, but it is less clear what factors may limit its acceptability and feasibility when combined with clinical care.

Objective: The purpose of this study was to explore barriers and facilitators to use of an adjunctive CCBT-I program among depressed patients in a psychiatric clinic by using both quantitative and qualitative approaches.

Methods: We conducted the qualitative component of the study using face-to-face or telephone interviews with participants who had enrolled in a clinical trial of a CCBT-I program as an adjunctive treatment in a psychiatric clinical setting. In line with the grounded theory approach, we used a semistructured interview guide with new thematic questions being formulated during the transcription and data analysis, as well as being added to the interview schedule. A range of open and closed questions addressing user experience were asked of all study participants who completed the 12-week trial in an online survey.

Results: Three themes emerged from the interviews and open questions, consistent with nonadjunctive CCBT-I implementation. Identification with the adjunctive intervention's target symptom of insomnia and the clinical setting were seen as key reasons to engage initially. Persistence was related to factors to do with the program, its structure, and its content, rather than any nonclinical factors. The survey results showed that only the key active behavioral intervention, sleep restriction, was rated as a major problem by more than 15% of the sample. In this clinical setting, the support of the clinician in completing the unsupported program was highlighted, as was the need for the program and clinical treatment to be coordinated.

Conclusions: The use of a normally unsupported CCBT-I program as an adjunctive treatment can be aided by the clinician's approach. A key behavioral component of the intervention, specific to insomnia treatment, was identified as a major problem for persistence. As such, clinicians need to be aware of when such components are delivered in the program and coordinate their care accordingly, if the use of the program is to be optimized.

ClinicalTrial: Australian and New Zealand Clinical Trials Registry ACTRN12612000985886; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=362875&isReview=true (Archived by WebCite at http://www.webcitation.org/6njjhl42X)

(J Med Internet Res 2017;19(2):e37) doi:10.2196/jmir.5639

KEYWORDS

adherence; persistence; eHealth; online intervention; Web-based intervention; motivations; barriers; insomnia; depression; men

Introduction

Depression is one of the most common psychiatric disorders in older adulthood and accounts for much of today's health care burden placed on medical professionals [1].

The recent development of eHealth technology has led to many efficacious interventions for depression being provided consistently to large numbers of people [2]. Computer-delivered therapies (often based on cognitive behavioral therapy [CBT] principles and thus termed computerized CBT [CCBT]) are more readily accessed by individuals who may regard formal mental health services as stigmatizing [3]. Typically, men are among the non-help-seeking group who may benefit from CCBT, as it is less costly and confrontational than face-to-face therapy, while still easily accessible via the growing utility of the Internet and smartphones [4]. When used correctly, CCBT can produce treatment effects similar to those observed in face-to-face treatment; however, those that have therapist interaction are shown to be more effective [2]. In their study, So et al [5] reported limited long-term gains from CCBT, along with an increased rate of participant withdrawal. This effect may be remedied by clinician program support [5]. A growing body of evidence shows that clinician-assisted CCBT results in significant improvements in patients with depressive [2,6] and anxiety disorders [2,7], with results comparable with those obtained from face-to-face treatment.

With rising health care costs, such CCBT programs have been proposed in clinical guidelines to help reduce the cost burden in primary and secondary health care. In 2006, the National Institute for Clinical Excellence [8] (now the National Institute for Health and Care Excellence), United Kingdom, included CCBT as first-step intervention in stepped-care treatment models. In these models, lower-cost interventions are tried first, with more-intensive and -costly interventions reserved for those insufficiently helped by the initial intervention [9]. Such models have proved to be cost effective for depression [10].

In reality, many such CCBT programs are used as adjunctive treatments to face-to-face or even telephone-delivered therapies, rather than with clinicians acting as adjuncts to the program. There is evidence that such adjunctive eHealth programs can add to the effectiveness of existing therapy. For example, Danaher et al [11] showed that combining a Web-based tobacco cessation intervention with using nicotine lozenges encouraged long-term tobacco and smokeless tobacco abstinence. There have been similar observations in other medical conditions, such as enhancing diabetes self-management as an adjunct to regular clinical support, and enhanced results of Internet toilet training for children with encopresis [12].

Given high levels of multimorbidity, and the heterogeneous nature of many psychiatric disorders, a further avenue in improving outcomes is the use of eHealth treatments for comorbidities or symptomatology of depression as adjunctive processes to face-to-face treatments. Insomnia is a frequently reported sleep disturbance in older adults with depression [1], and recent studies have shown that addressing insomnia by using CBT for insomnia can improve depression outcomes [13]. There are now efficacious computerized CBT for insomnia (CCBT-I) interventions [14], with evidence that they have some additional benefit on mood symptoms [15]. A recent study by Christensen et al [16] found that SHUTi significantly lowered depression symptoms on the Patient Health Questionnaire at 6 weeks and 6 months compared with HealthWatch, a general health psychoeducation program.

With the implementation of these adjunctive eHealth treatments into routine clinical practice, emphasis must be placed on process and other logistical concerns. It is less clear what factors may limit acceptability and feasibility of the effects seen in the Internet-only CCBT-I trials when used in clinical practice. Of particular interest are the ways in which patients are able to adhere to intervention techniques provided by the program with respect to timing, dose, and frequency (adherence), and the duration of time in which they are able to actively engage with therapy from initiation until discontinuation (persistence) [17]. Qualitative studies have implicated personal, design, and environmental factors associated with adherence and persistence to CCBT for both depression [18] and insomnia [19]. Adherence to Web-based programs [18] is related to outcome for depression. Further to this, clinicians making a referral or recommendation to a CCBT-I site enhance adherence to the program, suggesting therapy outcomes may relate to clinician interaction [20]. The purpose of this study was to explore the barriers and facilitators to the use of an adjunctive CCBT-I program among depressed patients alongside clinical care.

Methods

Study Sample

Participants in the study were primarily residents of New South Wales, Australia, who had enrolled in the Sleep Or Mood Novel Adjunctive therapy (SOMNA) trial, which included a CCBT-I program used as an adjunctive treatment in a psychiatric clinical setting [21]. All participants were English speaking, had an active email address, and currently met the criteria for at least minor depression, defined as screening positive on the Quick Inventory of Depressive Symptomatology [22]. A diagnosis of depression was then confirmed at a clinical interview by a psychiatrist, using *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) criteria for either a major depressive disorder or dysthymia, the Structured Clinical Interview for DSM-IV [23], and clinically significant insomnia symptoms reported on the Insomnia Severity Index [24]. Due to funding conditions, participants were male and aged 50 years or over. We excluded participants if they (1) had a history of psychosis or hypo/manic episode, (2) had a current substance dependence, or (3) had a score of <24 on the Mini-Mental State Examination, all as determined by the psychiatrist clinical interview; (4) were a rotating shift worker with overnight shifts, or with transmeridian travel >2 hours in the past month, (5) had all the criteria for restless legs syndrome as defined using the Cambridge-Hopkins Restless Legs Syndrome Short Form Questionnaire at screening, or (6) were at high risk for obstructive sleep apnea as defined by the Berlin Questionnaire or had been treated for obstructive sleep apnea. Participants were randomly allocated to either the active intervention of

CCBT-I program or to an active control of Internet-delivered sleep health information.

Clinical Setting

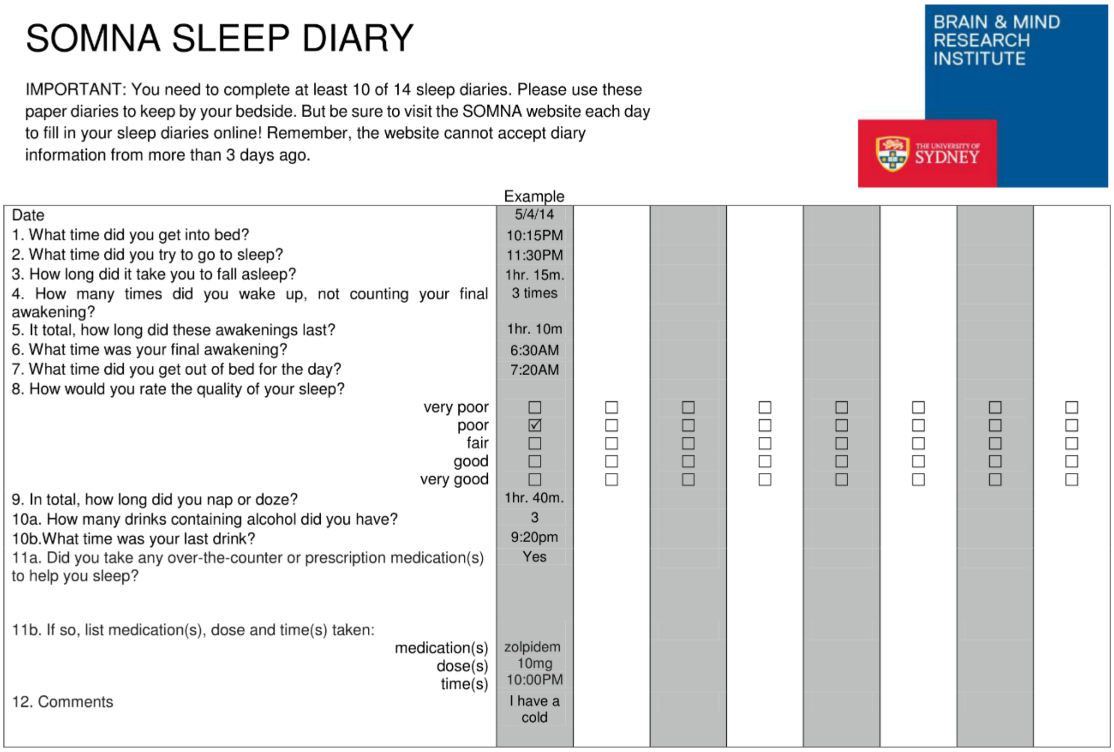
The participants were treated at a university psychiatric clinic staffed by consultant, final-year trainee psychiatrists, or both. Standard clinical care for treatment of depressive symptoms was delivered as considered clinically appropriate by the psychiatrist in conjunction with the *beyondblue* clinical guidelines [25]. This included watchful waiting, psychotherapy referral, or medication use under standard collaborative care with a general practitioner. As part of their standard treatment, participants were required to undertake a semistructured clinical assessment with a psychiatrist, whereupon a treatment plan was developed. This assessment included administration of standard clinical outcome measures, including the Hamilton Rating Scale for Depression, as well as completion of a full medical history. After 2 weeks, participants were required to undergo a clinical review and management appointment with the psychiatrist, in accordance with their treatment plan. This was reviewed at the first follow-up appointment at week 3. Over the subsequent 10 weeks, participants had further clinical review and management as clinically required. This management could also include the psychiatrist undertaking time-limited or supported therapies (although not focused on sleep problems), and referral to or liaison with psychologists, exercise physiologists, other allied health staff, and primary care. Standard clinical instruments were administered to monitor depressive symptoms over this period. If any changes to medication dosage were required, this would be discussed between the patient and psychiatrist or referring doctor.

CCBT-I Intervention: SHUTi

SHUTi (Sleep Health Using the Internet, BeHealth Solutions, LLC, Charlottesville, VA, USA) is a 9-week CCBT-I intervention. The program contains elements of psychoeducation, activity scheduling, thought challenging, problem solving, and interpersonal therapy. A major component of the program is sleep restriction, a standard behavioral strategy involving restricting the participant's time-in-bed (sleep window) to match their average self-reported total sleep duration. The sleep window is then titrated weekly based on sleep efficiency (the proportion of time-in-bed spent asleep), in order to arrive at the participant's core sleep requirement. Decreasing the opportunity to sleep over successive nights is postulated to build homeostatic sleep pressure, stabilize circadian control of sleep and wakefulness, and dampen presleep cognitive and physiological arousal.

The SHUTi program is fully automated with a new module opening each week after the last has been completed. It is organized into 6 weekly sessions, providing an overview in week 1, and focusing on sleep behavior in weeks 2 and 3, sleep thoughts in week 4, psychoeducation in week 5, and problem prevention in week 6. The participants were able to use the program over the 9 weeks to the end of the trial and then for a further 6 months if required. During the trial, participants received a reminder email 3-4 days after the module was opened or sleep diaries were due. Figure 1 is a sample of the sleep diary. The intervention was delivered alongside the clinical care, but the participants were repeatedly made aware that they could not communicate the components of their Internet program to the clinician to prevent loss of blinding.

Figure 1. Sample of the Sleep Or Mood Novel Adjunctive therapy (SOMNA) trial sleep diary.



SOMNA SLEEP DIARY

IMPORTANT: You need to complete at least 10 of 14 sleep diaries. Please use these paper diaries to keep by your bedside. But be sure to visit the SOMNA website each day to fill in your sleep diaries online! Remember, the website cannot accept diary information from more than 3 days ago.

	Example							
Date	5/4/14							
1. What time did you get into bed?	10:15PM							
2. What time did you try to go to sleep?	11:30PM							
3. How long did it take you to fall asleep?	1hr. 15m.							
4. How many times did you wake up, not counting your final awakening?	3 times							
5. It total, how long did these awakenings last?	1hr. 10m							
6. What time was your final awakening?	6:30AM							
7. What time did you get out of bed for the day?	7:20AM							
8. How would you rate the quality of your sleep?								
very poor	☐	☐	☐	☐	☐	☐	☐	☐
poor	☑	☐	☐	☐	☐	☐	☐	☐
fair	☐	☐	☐	☐	☐	☐	☐	☐
good	☐	☐	☐	☐	☐	☐	☐	☐
very good	☐	☐	☐	☐	☐	☐	☐	☐
9. In total, how long did you nap or doze?	1hr. 40m.							
10a. How many drinks containing alcohol did you have?	3							
10b.What time was your last drink?	9:20pm							
11a. Did you take any over-the-counter or prescription medication(s) to help you sleep?	Yes							
11b. If so, list medication(s), dose and time(s) taken:								
medication(s)	zolpidem							
dose(s)	10mg							
time(s)	10:00PM							
12. Comments	I have a cold							

Online Survey and Analysis

All participants who successfully completed the SOMNA trial were invited to take part in an online program evaluation survey immediately postintervention. The survey contained the Internet Intervention Adherence Questionnaire [26], together with some specific a priori questions relating to CCBT-I. This rates people's experiences and problems in adhering to the program on a 3-point scale. We grouped these into "major problem," "minor problem," and "no problem" and report results as n (%) for each. Within the survey, participants were presented with 2 open-ended items regarding their satisfaction with the program and their rating of the effectiveness of the SHUTi intervention in meeting their personal needs, as well as things such as the ability to offer constructive feedback about the overall user experience and accessibility of the program.

Semistructured Interviews and Analysis

Participants were recruited from those who had completed the final 6-month follow-up assessment (including the online survey above) of the SOMNA trial active arm by a research assistant independent of the SOMNA trial. We invited potential participants to the study by email and selected all interested respondents to participate. Following provision of a participant information statement and an overview of the study conduct, we asked potential participants to confirm participation as per written consent form, which stipulated that interviews would be audio recorded using an electronic recording device. We conducted recruitment and interviews in an iterative fashion, with transcripts being reviewed and responses collated, and continued until we reached theoretical saturation (where no new insights emerged from the data for 3 consecutive interviews). Saturation was achieved at the point of preliminary thematic analysis when participants did not report to the interviewer any information that constituted a discrete theme. This occurred after 7 interviews, and was confirmed by the interviewer at interview 10. Given the method and population targeted for recruitment, it was unlikely that subsequent interviews would produce novel or significantly disparate themes on analysis. We offered participants an interview time at their convenience, and we conducted interviews either face-to-face in a clinical setting or via telephone. Travel reimbursement was made available to face-to-face interviewees, although no participants required this.

Using a grounded theory approach, the chief investigator (NG) of the SOMNA trial developed a semistructured interview guide, allowing for new questions to be formulated and added to the interview schedule throughout data collection. Consequently, the interview schedule evolved over the course of the research project. The final interview guide included 10 primary items of interest: (1) overall experience of participation in the combined Internet and clinical interventions, (2) motivation to participate in SHUTi, (3) initial expectations about the Web-based intervention, (4) how expectations were met, (5) how expectations were not met, (6) motivation to continue with Web-based modules and diaries, (7) challenges while doing the modules, (8) experience of clinical treatment and Web-based program adjunctive delivery, (9) suggested improvements to SHUTi for general use, and (10) other reflections or comments on program user experience. Interviews were between 26 and 42 minutes in duration, with all interviews revolving around 1 key question: "What motivated you to persist with the [Web-based] SHUTi program?"

On completion of the interviews, an external agent transcribed the audio files, and the interviewer cross-checked the transcripts against the audio recordings. In conjunction with observational field notes, the chief investigator (NG) and coinvestigator (CC) then collaboratively coded the interview transcripts for recurrent response features. This was followed by focused thematic analysis, where we defined and pursued a central set of themes based on the prevalence and frequency of codes identified by the primary researcher. Conceptual memos were written from the focused themes to help develop an understanding of the themes and how these related to the data. We used no coding or analysis software for this analysis. We used method triangulation to check consistency.

Ethics Approval and Registration

We obtained written informed consent from all participants engaged with the SOMNA trial. Ethics approval for the SOMNA research trial and this qualitative substudy was obtained from the University of Sydney Human Research Ethics Committee, Sydney, Australia. The trial was registered on the Australian and New Zealand Clinical Trial Registry (ACTRN12612000985886).

Results

Sample

A total of 39 of the 40 (98%) eligible participants who commenced the SHUTi program and remained in the trial (regardless of their level of adherence to the modules) completed the survey. Participants were male and had an average age of 59 years. Of the 39 completers, 29 (74%) were married, 36 (92%) had completed post high school qualifications, 33 (85%) met DSM-5 criteria for a major depressive disorder, and 13 (33%) had self-reported cognitive complaints. Due to the anonymous nature of the interview protocol, we did not ascertain demographic data for the qualitative study participants.

Online Survey

Only one aspect of the program was reported as a major problem, and thus potentially a barrier to adherence, by more than 20% of the participants (Table 1). Difficulty with engaging with the sleep restriction component of the program was reported as a major problem by 12 of 39 respondents (31%). Approximately 1 in 9 participants reported major problems with external factors such as family issues (13%), work issues (10%), and no time (15%), and 13% of participants reported their sleep efficiency not improving. More than half (26/39, 67%) of participants found that CCBT-I improved the problem for which they sought help ("mostly," 11/39, 28%; "very," 15/39, 39%). Additionally, some participants (15/39, 39%) reported the program helped them feel prepared to handle this problem in the future ("most," 8/39, 21%; "very," 7/39, 18%).

Table 1. Self-report problems encountered during the Web-based SHUTi program (n=39).

Problems	Not a problem, n (%)	A little problem, n (%)	A major problem, n (%)	Not applicable, n (%)
Daily requirements were too much for me to do	34 (87)	2 (5)	3 (8)	0
Diaries were too long to complete	34 (87)	4 (10)	1 (3)	0
I had family issues	20 (51)	5 (13)	5 (13)	9 (23)
My family was on the computer	34 (87)	0	0	5 (13)
I had work issues	24 (62)	2 (5)	4 (10)	9 (23)
I forgot to enter data	21 (54)	11 (28)	3 (8)	4 (10)
I had a bad Internet connection	31 (80)	2 (5)	1 (3)	5 (13)
I had a slow Internet connection	29 (74)	3 (8)	1 (3)	6 (15)
I had no time	23 (59)	8 (21)	6 (15)	2 (5)
The screen was hard to read	37 (95)	2 (5)	0	0
Homework amount was too much	25 (64)	9 (23)	0	5 (13)
Homework was difficult	32 (82)	2 (5)	0	5 (13)
Sleep efficiency was not improving	14 (36)	16 (41)	5 (13)	4 (10)
Sleep restriction requirements were difficult	10 (26)	14 (36)	12 (31)	3 (8)
Too many sleep rules to follow	21 (54)	13 (33)	2 (5)	3 (8)
My spouse did not like the program	28 (72)	1 (3)	0	10 (26)
Too much text	28 (72)	10 (26)	1 (3)	0
Website had a bad connection	33 (85)	2 (5)	0	4 (10)
Website was difficult to understand	35 (90)	4 (10)	0	0
Website was hard to navigate	37 (95)	1 (3)	1 (3)	0
Website did not seem useful	32 (82)	2 (5)	5 (13)	0
Website was too long to do	29 (74)	9 (23)	1 (3)	0

Qualitative Study

We conducted a total of 4 telephone and 6 face-to-face interviews between September 2014 and February 2015.

Three main themes emerged from the qualitative interviews: (1) reasons to commence, and expectations prior to commencing the program related to intrinsic motivators, (2) motivators and barriers related to the program structure and content to continue with the program, and (3) clinical support. The content extracted from the face-to-face interviews regarding the motivators and barriers to continue with the program was triangulated with that collected from the self-report.

Reasons for Initial Engagement

Identification With the Adjunctive Component

Most participants reported they could identify themselves on the advertisement, which displayed the study inclusion criteria (men, age ≥50 years, with insomnia and stress). The advertisement was presented in various media forms (eg, magazines, online, poster) and was viewed by the participant prior to any involvement or enrollment with the SOMNA clinic. Participants reported being led by advertising material to approach the clinic in the first place, despite many already being in treatment for depression. Some stated they had ignored their insomnia for some time and the advertisement acted as a prompt to do something about it. Linked expectation included the hope that the program would improve their sleep. Some added they had hoped the program could also improve their mood and energy level (interviews 3 and 5). Surprisingly, some participants stated they had no real expectation of the program. Participant 7 said he had tried a few different methods but had not really found much success. He said he participated to “give it a go.” Most participants reported an intrinsic motivation to improve their health holistically. They understood that there was no “quick fix” for their insomnia, and they wanted to understand what the maintaining factors for chronic insomnia were.

Clinic Setting in a University

A total of 2 participants stated that they were interested in being involved in clinical research.

> *I try to be evidence based in my professional work, so I thought, well I will probably get a higher standard of care if I am on a trial than if I just go to some clinic or private practice.* [Interview 6]

This was echoed in perceptions that this setting might help where previous approaches had failed—for example, reporting having tried various types of remedies with little benefit. In some cases, this had led to some fears of entering the program, that “nothing works,” and even “anticipating failure” (interview 6).

Motivators to Persist With the Program

Design and Structure of the Program

Most participants reported they found the program easy to use and the contents easy to follow. Participant 1 said

> *I enjoy the visual representation…it feels good looking at the graphs…I look forward to putting them in to see how I went with getting my sleep efficiency and then felt good when I achieved that efficiency.*

Participant 2 agreed that the program fit well with the “visual person.” Participant 5 said that it suited his goal-directed personality and he found it rewarding to see the results.

Cognitive Component Through the Program Content

Most participants reported benefits from psychoeducation on sleep patterns, including understanding different factors that can influence sleep, developing good habits, and challenging cognitive misconceptions about sleep behaviors. Participants reported making “cognitive changes” and letting go of rigid reinforcing beliefs.

> *I learned from the program that it is okay to have minor setbacks…I was able to move on and try again the following night.*

Behavioral Component: Reinforcing Concepts Throughout the Program and Forming Habits

Participants reported that reinforcement of concepts throughout the modules made it easier to remember them when faced with challenges and insomnia. Most participants found completing a daily sleep diary helpful in making them aware of their sleeping pattern, and also allowing them to go back and review the patterns. This actual behavior and the feedback from the program became self-reinforcing during the program.

> *It becomes a habit for you so it was easier and that became the motivation, it was just something that you got up and did.* [Interview 5]

Some identified the behavioral focus early on in creating good habits.

> *The only way that it is going to be of any benefit to me is if I am regimented with, and do exactly, what they have asked me to do and just have faith in what was said, even though I am thinking how do I go to bed at one o’clock?* [Interview 1]
>
> *I think if in one area of your life you can become disciplined and regimented with it, that can help you to get disciplined in other areas of your life so it can have a flow-on effect.*

Impersonal and Immediately Accessible Nature of the Internet

Some of the participants preferred the impersonal nature of using the Web over going through the therapy with a psychologist.

> *The impersonal nature of using the Web was helpful for me. If I had to sit with a psychologist and go through the cores, the experience would be vastly different.*

The convenience and constant availability were also identified as the practical benefits of the Internet program in the postintervention feedback.

Barriers to Persisting

Design and Structure of the Program

On the self-reports, 2 participants suggested using newer graphics and videos. One participant, who worked in a graphic design background, stated “It could look more modern, as it is very conservative, but perhaps the current look is targeted to my demographic or older.” It was also suggested that the program could provide videos produced in Australia (the current videos are produced in the United States).

One participant found 1 of the modules to be too long. He said that “there was too much readings and going to next pages and back again. This would have been more beneficial in shorter modules.”

Some participants reported difficulties in completing the sleep diaries. One of them stated “the methodology to manually log sleep is crude and unreliable. The watch or other devices are more accurate method and should be integrated with the log.” In the self-reports, 3 participants commented that there should be more flexibility with making entries in the sleep diaries. One of them suggested a daily sleep chart question on daily health and whether it impacted on sleep. Some also reported a sense of failure if they did not complete the modules and diaries.

Technical Issues

Several recommendations were made regarding the technical aspects of the program. There were 2 participants who commented that the program should have increased compatibility with Apple iPad devices. One of them stated

> *Some sections within each of the cores would not appear on my iPad. The page would have the border etc but would be blank inside, next page and previous page would be fine. This technical issue needs fixing.*

Other suggestions include

> *A facility to retrieve or change one’s password is imperative. In addition to using dropdowns you should allow people to enter the information directly, such as typing minutes. When the user logs out the webpage is dimmed, which is not a good way to...*
>
> *A way to jump to any particular part/page directly or with a minimum key strokes.*
>
> *More printouts and more written information.*
>
> *Provide examples or interpretations of the intent behind various questions: eg, “how long did it take you to fall asleep?”—Is this during the sleep window (as the context suggests) or generally?*
>
> *Graphs only depicted in 1-week lots—it would be great to be able to nominate a time period or range*

> *of dates and see it all on one graph (therefore see actual trend line of improvement). More personal stories in video format please, need to see in time.*

Participant 10 had no access to a computer for a period during the treatment program so he could not respond to reminders, surveys, and emails.

Behavioral Component: Difficulties Adhering to Sleep Window

A key behavioral component of CBT for insomnia is sleep consolidation. This process involves using the data derived from people's sleep diaries to identify the optimum period and duration for their individual sleep. Counterintuitively, this usually leads to a period of sleep restriction. Most participants found it challenging to stick with the sleep window, in particular finding it difficult to stay up until the designated sleep time. For example: "I would sit there and I was in no-man's-land." Some were concerned about the effects of the fatigue due to the initial small sleep window on their ability to function and work the following day. Some suggested users should be advised to identify a window of time for their period of sleep deprivation. It was also difficult to stay awake well after the bedtime of their partners, something that integration with the clinical aspects may have enhanced.

The Use of CCBT-I in a Clinical Setting

Due to the nature of the double-blind randomized trial, clinicians were unaware whether a participant was allocated to the active computer program intervention or the generic psychoeducation program. It was therefore not possible to integrate the clinical care fully into the computer program intervention in this study.

Nevertheless, most participants found the comprehensive psychiatric assessment and management helpful both for themselves and in persisting with the program: "I would not have completed the program without the clinical support" (interview 3). Conversely, 1 participant perceived the clinical setting as a potential barrier because of either the stigma associated with psychiatry (interview 5) or the logistical difficulties of attending a clinic during working hours.

Two participants (2 and 10) commented that the clinical and computer program interventions and advice seemed disjointed at times. One commented that there may have been a disconnect between the computer results and the clinical assessment, and the clinicians could have been more in tune to the participants' input in the Web-based program in regard to sleeping patterns (interview 2).

Discussion

This study aimed to identify the factors influencing adherence to adjunctive CCBT-I in a clinical setting delivering treatment for depression and insomnia. Three themes emerged from the interviews and self-report questionnaires: (1) reasons to commence, and expectations prior to commencing the program, which related to intrinsic motivators, (2) motivators and barriers to continue with the program in relation to the program structure and content, and (3) the interaction of the program with clinical management.

Reasons to Participate in and Commence the Program

Most of the participants who underwent qualitative interviews about the SOMNA trial stated they could identify themselves in the advertisement. Some had previously been unsuccessfully treated for insomnia and depression. Some reported an intrinsic motivation to better their health holistically, which helped this group of participants persist with the program. The belief in recovery and helpfulness of the program as suggested by Barazzone et al [27] appeared to foster establishment of a therapeutic relationship with the program.

Some participants were attracted to joining the program, as it was conducted within a clinical trial setting, despite seemingly having little expectation or incentive. The reasons, of perceived better quality of care in a university trial setting, were slightly different from the more altruistic reasons for engaging in a trial, that it would benefit others, reported by Donkin and Glozier [18].

Motivators and Barriers to Continue With the Program

Program Structure

It appeared that the program structure of SHUTi had facilitated the participants' adhering to the Web-based intervention. A majority of participants reported they derived clinical improvements from SHUTi in having a structure to follow. Evidence showed that improving therapeutic relationships would foster a working alliance, which might be associated with proactive information seeking and adherence to eHealth programs [28]. Our participants reported that visual representations had helped them to understand the concept of sleep efficiency and how to improve it. This is in line with the findings that information visualization could be an effective supportive application to improve people's health literacy [29].

Cognitive Component: Psychoeducation Through the Program Content

In our study, psychoeducation (on insomnia) appeared to be regarded well by the participants. Many of the participants reported gaining more knowledge during the program. This was particularly true for those who had ignored their insomnia as a problem for a long time and had little or no previous interventions for their insomnia. Psychoeducation corrected the participants' misconceptions about sleep behaviors. The results from the survey showed that more than half of the participants found the website improved their insomnia and helped them feel prepared to handle symptoms of insomnia in the future.

Behavioral Component: Reinforcing Concepts Throughout the Program and Forming Habits

The visual presentation in SHUTi, such as the graphic interpretations of sleep efficiency, was cited as particularly beneficial for those who had a goal-directed personality. The videos appeared to help participants to relate to others with similar sleep problems. Key concepts were reinforced throughout the modules, and this helped the participants in forming habits, making it easier to persist with the program.

The data entry and self-evaluation of sleep efficiency enabled the participants to review their sleep patterns and receive tailored feedback on their progress. The Web-based program also allowed the participants to learn and practice at their own pace. In this way, these appear to provide feedback, responsiveness, and flexibility, as identified by Barazzone et al [27] as important elements in developing and maintaining a therapeutic relationship.

In terms of barriers identified in this trial, aspects of CBT for insomnia in general and program-specific technical issues in particular could have hindered the participants' adherence to the Web-based program. For instance, many participants highlighted the intrinsic difficulty of sleep restriction as challenging. These challenges are present in face-to-face CBT for insomnia too. In their exploratory study on patients engaging in sleep restriction therapy for insomnia, Kyle et al [30] reported that one-third of the participants mentioned impaired driving ability due to sleep restriction, and several reported side effects (such as headache, nausea) during the first week of the restriction.

Some participants found the modules were too lengthy. The modules in SHUTi typically involve up to an hour of engagement, over a 9-week period [21]. Previous authors [31] have reported that CCBT programs are typically accessed for much briefer periods than the traditional therapy session duration of 50 minutes. As such, disaggregating the modules into shorter pieces might have helped engagement, which is key to an effective intervention [18].

The Use of the CCBT-I in a Clinical Setting

Evidence from eHealth literature suggested that Web-based intervention might reduce stigma and allow more people seek help for mental health conditions [3]. In our sample, a small proportion of participants preferred the impersonal nature of using the Web over going through the cores with a clinician. One participant identified clinical care as a potential barrier due to the stigma of requiring psychiatric treatment. This might be due to one's perception of mental health and cultural background, which influence their view on seeking psychiatric care.

In the qualitative data, the participants did not think the website helped reduce their contact with health professionals. Most participants reported that the regular clinical reviews facilitated their adherence to the Web-based program. This group said the clinical reviews provided them with a sense of accountability to complete the online tasks. They found that the reviews promoted their intrinsic motivation, and some added that the recommended clinical interventions, such as medications and mindfulness techniques, helped them to persist with the Web-based program. This may be why we did not identify some factors that emerged in nonclinical settings that were linked to persistence, such as a sense of "getting things done" and satisfaction at completing the program, as these may have been subsumed by the clinic.

The clinical management appeared to aid the CCBT-I by setting goals, planning treatment, and selecting intervention; that is, providing individualized treatment to the individual. This is essential in CBT, and indeed in any treatment. These elements would likely be more effectively achieved by incorporating a clinician into the Web-based intervention. On the other hand, the impersonal nature and constant availability of the Web-based program made attending the program more convenient for participants.

As Barazzone et al [27] suggested, CCBT-I programs encompass a substantial body of features essential to establishing a therapeutic relationship, but fewer features designed to develop and maintain a therapeutic alliance. Adjunctive clinical reviews could fill this void, by holding users through difficult phases of therapy, setbacks, and ruptures [32]. This may be a factor in the greater effectiveness of supported CCBT-I, as identified in the meta-analysis by Zachariae et al [14], and the clinical context of using unsupported CBT as an adjunctive treatment may lead to great benefits, although this comparison is yet to be tested.

Further, while most participants found the clinical intervention helpful, others found the clinical and computer program disjointed. This lack of coordination could be explained by the nature of the double-blind randomized trial. However, it highlights that in future commercially available CCBT-I programs need to have some facility to provide treating clinicians with information on where the participant is in the program, what interventions and messages are being delivered at that time, and any clinical progress. Otherwise, clinicians will be unable to integrate clinical intervention with the specific contents in the Web-based programs and potentially disengage from the programs. Clinicians in this study were therefore not aware of which intervention the participants received and were encouraged not to ask specific questions about the Web-based interventions in order to avoid unblinding.

Limitations

There are several limitations to our study. Our results are unique to the targeted population (ie, men aged ≥50 years) and the analysis of this research team. Consideration should be given to the variance in results that may have been produced by focusing on a less-specific cohort. For example, studying women or a younger adult population could conceivably yield different findings. However, for the purposes of this study, this is of little consequence to the anticipated outcomes. The data analysis was also influenced by the perspective of the researchers. In this study, the primary author worked as a clinical psychiatrist. This position could have resulted in the development of clinically biased views of what influences adherence behavior in a clinical setting. This, however, may be different from the influences of adherence in a research or online forum. Given this, we attempted to be as objective as possible, but the analysis is likely to have been couched within these experiences.

The participants interviewed were people who volunteered to participate in a qualitative interview from a group that volunteered to participate in a trial. This group was therefore probably highly motivated to participate in research and may be different from those who do not wish to share their views. This is reflected in the reported emphasis of research as an important motivator by some interview participants.

Due to trial blinding constraints, we contacted participants over a period of months, some participants not until approximately 6 months after they had completed the intervention. This may have affected the ability of some participants to recall their intrinsic motivations for continuing to engage with the trial, and their recall is likely to have been influenced by their present situation. While some of the themes that emerged appear to be values based, and therefore quite static, it is likely that state variables such as time factors and frustration may not be as accurately reflected in these findings. Also, in the self-report data, the participants were not asked about what factors motivated their ongoing participation.

We conducted 4 of the 10 interviews over the telephone. While phone interviewing may have allowed respondents to feel relaxed and better able to disclose sensitive information, the absence of visual cues may have resulted in loss of contextual and nonverbal impressions on the interviewer, and may have inadvertently compromised rapport, probing, and interpretation of responses. By using a semistructured, standardized interview schedule, we attempted to consistently present interview questions in the same arrangement and verbal manner with all participants, while simultaneously encouraging elaboration of responses through directive provocation and unlimited time for responses. Further to this, the interviewer adhered to neutral responses (eg, “okay,” “go on”) to prevent leading the participant. Any environmental bias was minimized by the implementation of standard interview practice requiring a quiet, comfortable room in which to conduct the interview, and was moderated by any specific requests of the interviewee (eg, specific time of day).

Finally, we were only able to recruit participants who had completed the trial, although many had not necessarily completed all the modules; that is, they persisted with the trial but were not completely adherent to the intervention, or were withdrawn, preventing us from exploring the barriers and motivators in people who did not persist with the intervention. We were also not able to recruit anyone who completed the Web-based program but not the 6-month follow-up assessment. Of those who did not persist and did provide additional contact with the researchers, their reasons for dropping out of the trial tended to comprise of a lack of time to complete the program at present, and therefore an unwillingness to participate in a further interview. This, together with other limitations discussed above, might have contributed to an early saturation in our data collection after 10 interviews. This is a relatively small sample size in similar qualitative research. Future research would benefit from interviewing participants who have dropped out of Web-based interventions to determine their reasons for doing so.

Conclusion

In terms of maintaining participants in Web-based programs, it may be useful to consider ways to maximize engagement and overcome the barriers that participants mentioned. The findings from this study indicated that the Web-based intervention benefitted those who were motivated, had a goal-directed personality, liked graphic interpretations, and were interested in improving their health holistically. The participants reported a benefit from the psychoeducation and behavioral reinforcement throughout the program in terms of engagement. Clinicians may capitalize on the motivation of this group of participants by providing regular feedback on the data provided and working on goals; for example, designing sleep restriction windows may enhance engagement and increase the likelihood of completion.

Other findings from this study suggested that care needs to be taken to ensure that messages and interventions given by the program and the clinician are coordinated if the use of the program is to be optimized. Future studies could further explore ways of integrating clinical input to enhance adherence and optimize treatment. For instance, Cavanagh [33] suggested that clinicians acknowledge both the strengths and limitations of CCBT-I programs, and offer them in the context of appropriate support that starts with discussion around the motivation, expectations, and preconceptions of the user regarding CCBT-I, addresses concerns, and promotes ongoing engagement with the program.

Clinicians may need to provide more additional support for those participants who are less motivated and goal directed. It would give these individuals realistic expectations and minimize anticipatory failure. Some participants mentioned online forums to promote social support and share knowledge. This may enhance a sense of community, and thus increase the likelihood of completing the program. Such forums have been successfully used in other programs [34]. Comparable with real-life clinical settings, there were participants in this study who had tried various interventions for their insomnia with little perceived benefit. This group might benefit from a more assertive and integrated approach in clinical and Web-based interventions.

Integrating automated eHealth interventions into routine clinical practice may improve outcomes; however, the programs and clinicians need to work in a synergistic fashion, sharing information and supporting the patient in their engagement with both aspects of treatment, if we are to realize the full potential of such approaches.

Conflicts of Interest

None declared.

References

1. Cho HJ, Lavretsky H, Olmstead R, Levin MJ, Oxman MN, Irwin MR. Sleep disturbance and depression recurrence in community-dwelling older adults: a prospective study. Am J Psychiatry 2008 Dec;165(12):1543-1550 [FREE Full text] [doi: 10.1176/appi.ajp.2008.07121882] [Medline: 18765482]

2. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
3. Griffiths KM, Carron-Arthur B, Parsons A, Reid R. Effectiveness of programs for reducing the stigma associated with mental disorders. A meta-analysis of randomized controlled trials. World Psychiatry 2014 Jun;13(2):161-175 [FREE Full text] [doi: 10.1002/wps.20129] [Medline: 24890069]
4. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
5. So M, Yamaguchi S, Hashimoto S, Sado M, Furukawa TA, McCrone P. Is computerised CBT really helpful for adult depression?-A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity. BMC Psychiatry 2013;13:113 [FREE Full text] [doi: 10.1186/1471-244X-13-113] [Medline: 23587347]
6. Vernmark K, Lenndin J, Bjärehed J, Carlsson M, Karlsson J, Oberg J, et al. Internet administered guided self-help versus individualized e-mail therapy: a randomized trial of two versions of CBT for major depression. Behav Res Ther 2010 May;48(5):368-376. [doi: 10.1016/j.brat.2010.01.005] [Medline: 20152960]
7. Titov N, Andrews G, Robinson E, Schwencke G, Johnston L, Solley K, et al. Clinician-assisted Internet-based treatment is effective for generalized anxiety disorder: randomized controlled trial. Aust N Z J Psychiatry 2009;43(10):905-912 [FREE Full text] [doi: 10.1080/00048670903179269]
8. National Institute for Health and Care Excellence. Computerised cognitive behaviour therapy for depression and anxiety (TA97). Technology Appraisal Guidance. London, UK: NICE; 2006 Feb 22. URL: https://www.nice.org.uk/guidance/ta97/resources/computerised-cognitive-behaviour-therapy-for-depression-and-anxiety-2294884489669 [accessed 2017-01-23] [WebCite Cache ID 6njhnhmzv]
9. Sobell MB, Sobell LC. Stepped care as a heuristic approach to the treatment of alcohol problems. J Consult Clin Psychol 2000 Aug;68(4):573-579. [Medline: 10965632]
10. van Straten A, Seekles W, van 't Veer-Tazelaar NJ, Beekman ATF, Cuijpers P. Stepped care for depression in primary care: what should be offered and how? Med J Aust 2010 Jun 7;192(11 Suppl):S36-S39. [Medline: 20528706]
11. Danaher B, Severson H, Crowley R, van Meter N, Tyler M, Widdop C, et al. Randomized controlled trial examining the adjunctive use of nicotine lozenges with MyLastDip: an eHealth smokeless tobacco cessation intervention. Internet Intervent 2015;2(1):69-76.
12. Ritterband LM, Cox DJ, Walker LS, Kovatchev B, McKnight L, Patel K, et al. An Internet intervention as adjunctive therapy for pediatric encopresis. J Consult Clin Psychol 2003 Oct;71(5):910-917. [doi: 10.1037/0022-006X.71.5.910] [Medline: 14516239]
13. Manber R, Edinger JD, Gress JL, San Pedro-Salcedo MG, Kuo TF, Kalista T. Cognitive behavioral therapy for insomnia enhances depression outcome in patients with comorbid major depressive disorder and insomnia. Sleep 2008 Apr;31(4):489-495 [FREE Full text] [Medline: 18457236]
14. Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia - A systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev 2016 Dec;30:1-10. [doi: 10.1016/j.smrv.2015.10.004] [Medline: 26615572]
15. Ye Y, Zhang Y, Chen J, Liu J, Li X, Liu Y, et al. Internet-based cognitive behavioral therapy for insomnia (ICBT-i) improves comorbid anxiety and depression: a meta-analysis of randomized controlled trials. PLoS One 2015;10(11):e0142258 [FREE Full text] [doi: 10.1371/journal.pone.0142258] [Medline: 26581107]
16. Christensen H, Batterham PJ, Gosling JA, Ritterband LM, Griffiths KM, Thorndike FP, et al. Effectiveness of an online insomnia program (SHUTi) for prevention of depressive episodes (the GoodNight Study): a randomised controlled trial. Lancet Psychiatry 2016 Jan 27;3(4):333-341. [doi: 10.1016/S2215-0366(15)00536-2] [Medline: 26827250]
17. Cramer JA, Roy A, Burrell A, Fairchild CJ, Fuldeore MJ, Ollendorf DA, et al. Medication compliance and persistence: terminology and definitions. Value Health 2008;11(1):44-47 [FREE Full text] [doi: 10.1111/j.1524-4733.2007.00213.x] [Medline: 18237359]
18. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. J Med Internet Res 2012;14(3):e91 [FREE Full text] [doi: 10.2196/jmir.2100] [Medline: 22743581]
19. Middlemass J, Davy Z, Cavanagh K, Linehan C, Morgan K, Lawson S, et al. Integrating online communities and social networks with computerised treatment for insomnia: a qualitative study. Br J Gen Pract 2012 Dec;62(605):e840-e850 [FREE Full text] [doi: 10.3399/bjgp12X659321] [Medline: 23211265]
20. Batterham PJ, Neil AL, Bennett K, Griffiths KM, Christensen H. Predictors of adherence among community users of a cognitive behavior therapy website. Patient Prefer Adherence 2008;2:97-105 [FREE Full text] [Medline: 19920949]
21. Cockayne NL, Christensen HM, Griffiths KM, Naismith SL, Hickie IB, Thorndike FP, et al. The Sleep Or Mood Novel Adjunctive therapy (SOMNA) trial: a study protocol for a randomised controlled trial evaluating an internet-delivered cognitive behavioural therapy program for insomnia on outcomes of standard treatment for depression in men. BMC Psychiatry 2015 Feb 05;15:16 [FREE Full text] [doi: 10.1186/s12888-015-0397-x] [Medline: 25652579]

22. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 01;54(5):573-583. [Medline: 12946886]
23. Spitzer R. SCID 101 for DSM-IV: Training Video for the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID). New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 1996.
24. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [Medline: 11438246]
25. Ellis P, Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression. Australian and New Zealand clinical practice guidelines for the treatment of depression. Aust N Z J Psychiatry 2004 Jun;38(6):389-407. [doi: 10.1111/j.1440-1614.2004.01377.x] [Medline: 15209830]
26. Ritterband LM, Ardalan K, Thorndike FP, Magee JC, Saylor DK, Cox DJ, et al. Real world use of an Internet intervention for pediatric encopresis. J Med Internet Res 2008 Jun 30;10(2):e16 [FREE Full text] [doi: 10.2196/jmir.1081] [Medline: 18653440]
27. 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]
28. 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]
29. Faisal S, Blandford A, Potts HW. Making sense of personal health information: challenges for information visualization. Health Informatics J 2013 Sep;19(3):198-217. [doi: 10.1177/1460458212465213] [Medline: 23981395]
30. Kyle SD, Morgan K, Spiegelhalder K, Espie CA. No pain, no gain: an exploratory within-subjects mixed-methods evaluation of the patient experience of sleep restriction therapy (SRT) for insomnia. Sleep Med 2011 Sep;12(8):735-747. [doi: 10.1016/j.sleep.2011.03.016] [Medline: 21907616]
31. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. J Med Internet Res 2002;4(1):e3 [FREE Full text] [doi: 10.2196/jmir.4.1.e3] [Medline: 11956035]
32. Richardson R, Richards DA. Self-help: towards the next generation. Behav Cogn Psychother 2006;34(01):13-23. [doi: 10.1017/S1352465805002481]
33. Cavanagh K. Turn on, tune in and (don't) drop out: engagement, adherence, attrition and alliance with Internet-based CBT interventions. In: Bennett-Levy J, Richards D, Farrand P, Christensen H, Griffiths K, Kavanagh D, et al, editors. Oxford Guide to Low Intensity CBT Interventions. Oxford, UK: Oxford University Press; 2010:227-233.
34. Espie CA, Kyle SD, Williams C, Ong JC, Douglas NJ, Hames P, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep 2012 Jun;35(6):769-781 [FREE Full text] [doi: 10.5665/sleep.1872] [Medline: 22654196]

Abbreviations

CBT: cognitive behavioral therapy
CCBT: computerized cognitive behavioral therapy
CCBT-I: computerized cognitive behavioral therapy for insomnia
DSM-5: Diagnostic and Statistical Manual of Mental Disorders
SOMNA: Sleep Or Mood Novel Adjunctive therapy

Edited by G Eysenbach; submitted 12.02.16; peer-reviewed by H Potts, Q Yuan; comments to author 10.03.16; revised version received 03.06.16; accepted 21.01.17; published 10.02.17.

Please cite as:
Chan C, West S, Glozier N
Commencing and Persisting With a Web-Based Cognitive Behavioral Intervention for Insomnia: A Qualitative Study of Treatment Completers
J Med Internet Res 2017;19(2):e37
URL: http://www.jmir.org/2017/2/e37/
doi:10.2196/jmir.5639
PMID:28188124

Original Paper

An Online Intervention Comparing a Very Low-Carbohydrate Ketogenic Diet and Lifestyle Recommendations Versus a Plate Method Diet in Overweight Individuals With Type 2 Diabetes: A Randomized Controlled Trial

Laura R Saslow[1], PhD; Ashley E Mason[2], PhD; Sarah Kim[3], MD; Veronica Goldman[2]; Robert Ploutz-Snyder[4], PhD; Hovig Bayandorian[5]; Jennifer Daubenmier[6], PhD; Frederick M Hecht[2*], MD; Judith T Moskowitz[7*], PhD

[1]Department of Health Behavior and Biological Sciences, School of Nursing, University of Michigan, Ann Arbor, MI, United States
[2]Osher Center for Integrative Medicine, School of Medicine, University of California, San Francisco, San Francisco, CA, United States
[3]School of Medicine, University of California, San Francisco, San Francisco, CA, United States
[4]Applied Biostatistics Laboratory, School of Nursing, University of Michigan, Ann Arbor, MI, United States
[5]University of California, Berkeley, Berkeley, CA, United States
[6]Institute of Holistic Health, Department of Health Education, San Francisco State University, San Francisco, CA, United States
[7]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States
[*]these authors contributed equally

Corresponding Author:
Laura R Saslow, PhD
Department of Health Behavior and Biological Sciences
School of Nursing
University of Michigan
Room 2178
400 N Ingalls St
Ann Arbor, MI, 48105
United States
Phone: 1 734 764 7836
Fax: 1 734 936 5525
Email: saslowl@umich.edu

Related Articles:

Comment in: http://www.jmir.org/2018/5/e180/

Comment in: http://www.jmir.org/2018/5/e181/

Abstract

Background: Type 2 diabetes is a prevalent, chronic disease for which diet is an integral aspect of treatment. In our previous trial, we found that recommendations to follow a very low-carbohydrate ketogenic diet and to change lifestyle factors (physical activity, sleep, positive affect, mindfulness) helped overweight people with type 2 diabetes or prediabetes improve glycemic control and lose weight. This was an in-person intervention, which could be a barrier for people without the time, flexibility, transportation, social support, and/or financial resources to attend.

Objective: The aim was to determine whether an online intervention based on our previous recommendations (an ad libitum very low-carbohydrate ketogenic diet with lifestyle factors; “intervention”) or an online diet program based on the American Diabetes Associations’ “Create Your Plate” diet (“control”) would improve glycemic control and other health outcomes among overweight individuals with type 2 diabetes.

Methods: In this pilot feasibility study, we randomized overweight adults (body mass index ≥25) with type 2 diabetes (glycated hemoglobin [HbA_{1c}] 6.5%-9.0%) to a 32-week online intervention based on our previous recommendations (n=12) or an online diet program based around a plate method diet (n=13) to assess the impact of each intervention on glycemic control and other health outcomes. Primary and secondary outcomes were analyzed by mixed-effects linear regression to compare outcomes by group.

Results: At 32 weeks, participants in the intervention group reduced their HbA_{1c} levels more (estimated marginal mean [EMM] –0.8%, 95% CI –1.1% to –0.6%) than participants in the control group (EMM –0.3%, 95% CI –0.6% to 0.0%; *P*=.002). More than half of the participants in the intervention group (6/11, 55%) lowered their HbA_{1c} to less than 6.5% versus 0% (0/8) in the control group (*P*=.02). Participants in the intervention group lost more weight (EMM –12.7 kg, 95% CI –16.1 to –9.2 kg) than participants in the control group (EMM –3.0 kg, 95% CI –7.3 to 1.3 kg; *P*<.001). A greater percentage of participants lost at least 5% of their body weight in the intervention (10/11, 90%) versus the control group (2/8, 29%; *P*=.01). Participants in the intervention group lowered their triglyceride levels (EMM –60.1 mg/dL, 95% CI –91.3 to –28.9 mg/dL) more than participants in the control group (EMM –6.2 mg/dL, 95% CI –46.0 to 33.6 mg/dL; *P*=.01). Dropout was 8% (1/12) and 46% (6/13) for the intervention and control groups, respectively (*P*=.07).

Conclusions: Individuals with type 2 diabetes improved their glycemic control and lost more weight after being randomized to a very low-carbohydrate ketogenic diet and lifestyle online program rather than a conventional, low-fat diabetes diet online program. Thus, the online delivery of these very low-carbohydrate ketogenic diet and lifestyle recommendations may allow them to have a wider reach in the successful self-management of type 2 diabetes.

Trial Registration: ClinicalTrials.gov NCT01967992; https://clinicaltrials.gov/ct2/show/NCT01967992 (Archived by WebCite at http://www.webcitation.org/6o0fI9Mkq)

(J Med Internet Res 2017;19(2):e36) doi:10.2196/jmir.5806

KEYWORDS

eHealth; diet; weight loss; type 2 diabetes mellitus

Introduction

Type 2 diabetes mellitus is a rapidly growing chronic disease that affects approximately 22 million people in the United States, for which diet is an integral aspect of treatment [1,2]. Data suggest that very low-carbohydrate diets [3-11], and adequate sleep and physical exercise [12-16] can improve glycemic control and reduce body weight in individuals with type 2 diabetes. Moreover, behavioral adherence strategies, including positive affect regulation and mindful eating strategies, may reduce overall stress, stress-based eating, and depressive symptoms, which can be barriers for following behavioral recommendations [17-19].

In prior research, we found that recommendations to follow a very low-carbohydrate diet and to make lifestyle changes (sleep and exercise recommendations and a package of behavioral adherence strategies based on positive affect regulation and mindfulness) were able to improve glycemic control and reduce body weight in overweight individuals with type 2 diabetes or prediabetes [20]. Although promising, this previous trial was delivered in-person, which is a significant barrier to engagement for people without the time, flexibility, transportation, social support, and/or financial resources to attend. To create a highly disseminable, evidence-based program, we adapted our in-person intervention for online delivery.

In this pilot feasibility and acceptability study, we assessed whether overweight individuals with type 2 diabetes, randomized to receive an online intervention based on our previous trial (recommendations to follow an ad libitum very low-carbohydrate ketogenic diet and other lifestyle changes), would have greater improvements in glycemic control and other health outcomes than participants randomized to a control group, an online diet program based on a plate method diet (the American Diabetes Associations' "Create Your Plate" diet). To our knowledge, this is the first online randomized controlled trial to teach a very low-carbohydrate ketogenic diet to individuals with type 2 diabetes.

Methods

Participants and Procedure

We conducted a parallel-group, balanced randomization (1:1) trial, approved by the University of California, San Francisco, Institutional Review Board and registered with ClinicalTrials.gov (NCT01967992). The primary outcome measure was glycemic control, operationalized as change in glycated hemoglobin (HbA_{1c}). A key secondary outcome was body weight. Exploratory outcomes were cholesterol, triglycerides, diabetes-related distress, subjective experiences of the diet, and physical side effects.

We recruited participants nationally with online ads (on Craigslist, Backpage, and Facebook), newspaper ads and articles, and radio ads. This allowed us to recruit participants from across the United States. Eligibility criteria included age 18 years or older with a body mass index of ≥25, an elevated HbA_{1c} level diagnostic of type 2 diabetes (6.5%-9%, measured by us at baseline), and regular access to the Internet. To reduce the risk of hypoglycemia, we excluded participants who were taking any diabetes medication other than metformin.

Recruitment materials directed interested participants to a study website to complete an online eligibility questionnaire. Study staff then called potentially eligible participants to assess initial eligibility and describe study procedures. For example, we assessed whether participants were taking any medications for their type 2 diabetes other than metformin. If participants then consented to the full study, they were asked to complete several assessments, all specified in the measures section.

We recruited participants who were ready to make the changes required of the intervention in order to mitigate a potentially high dropout rate. We measured readiness to undertake the intervention [21] with the following item: “If you are eligible for this study you will be asked to...cut out the kinds of cookies, cakes, pasta, pastries, bagels, rice, potatoes, and sugary fruits that some people often eat. Given the description of the dietary changes above, how prepared do you feel to make these changes?” Participants could answer the item from 1 (not at all) to 7 (very much so). Participants were eligible if they rated themselves to be prepared to begin above the midpoint of the scale (5-7). In addition, conscientious people, we reasoned, would be more likely to follow the behavior changes requested by the intervention because conscientiousness has been shown to be positively related to following beneficial health-related behaviors [22]. Participants were eligible if they rated themselves as conscientious on two items [23]: “I see myself as someone who is dependable, self-disciplined” (eligible answers were agree and strongly agree) and “I see myself as someone who is disorganized, careless” (eligible answers were disagree and strongly disagree). We used items from the Yale Food Addiction Scale [24] and the Eating Disorder Diagnostic Scale [25] to screen out participants who had the tendency to be addicted to food or binge eat.

For this study, it was not possible for the participants and staff to be masked to group allocation. Therefore, once all baseline measurements had been completed, study staff randomized participants to one of the two intervention groups by opening the next opaque envelope in a series containing the concealed sequence for randomization, which was created by a statistician using block randomization procedures, with blocks of size randomly allocated to size 2, 4, or 6.

Outcomes were measured at baseline as well as 16 and 32 weeks after baseline. We paid participants US $25 for each assessment at 16 and 32 weeks. Thus, participants could receive up to US $50 over the course of the study.

Intervention

Intervention Group: Very Low-Carbohydrate Ketogenic Diet and Lifestyle Recommendations

We randomized half of participants to receive recommendations on how to eat an ad libitum very low-carbohydrate ketogenic diet, to reduce carbohydrate intake to between 20-50 grams of nonfiber carbohydrates a day with the goal of restricting carbohydrate intake to a level at which a low amount of ketone production is induced, called nutritional ketosis. In this state, the body uses fatty acids instead of carbohydrates as its primary energy source, which do not elevate glucose levels as strongly as carbohydrates [26,27]. To support dietary adherence, we mailed participants in this group urinary acetoacetate (a ketone that can be measured in urine) test kits (KetoStix, Abbott). We asked them to measure their urine for the presence of ketones at least once a week for the first few months of the program.

We also suggested that participants in the intervention group follow lifestyle recommendations, including behavioral adherence strategies aimed at increasing positive affect regulation [18] and mindful eating based largely on the Mindfulness-Based Eating Awareness Training program [19,28], using handouts and lesson content adapted from our in-person intervention. Specific topics included setting attainable goals; scheduling, noticing, and savoring positive events; developing self-compassion; practicing positive reappraisal, gratitude, and acts of kindness; being aware of one’s personal strengths; and being mindful of hunger, fullness, cravings, taste satisfaction, and triggers for overeating. Moreover, starting in week 6, the lessons discussed the importance of physical activity and sleep as well as encouraged participants to increase their level of physical activity and amount of sleep. We chose to include a comprehensive program of behavioral support in this intervention group in order to enhance the likelihood of finding an impact of our previously successful in-person program using an online format.

We emailed participants in this group new lessons weekly for the first 16 weeks and then every two weeks for the remaining 16 weeks of the study. The lessons in the first 16 weeks included short videos created for the study about all of the study components (about 5-15 minutes long, with audio narration over videos with white text, images, and a black background), with printable handouts and links to online resources, such as recipes and recipe books. The lessons in the last 16 weeks did not include study-specific videos, only printable handouts and links to online resources.

Control Group: American Diabetes Associations’ “Create Your Plate” Diet

This dietary intervention, our control group, was slightly different from the one we had originally used in our in-person intervention because we received feedback that the previous “carbohydrate counting” intervention was difficult for participants to follow. Instead, we randomized half of participants to a diet program based around a plate method diet, the American Diabetes Associations’ “Create Your Plate” diet, a low-fat diet that emphasizes green vegetables, lean protein sources, and somewhat limited starchy and sweet foods. All proportions are based on a 9-inch plate: half the plate is filled with nonstarchy vegetables, one-quarter with carbohydrates, and one-quarter with lean proteins [29]. We taught this group using short videos created for the study (approximately 5-10 minutes long), with printable handouts and links to online resources, such as links to online recipes and recipe books. We chose to include just the standard dietary information in this group, and not all the extra behavioral help, in order to have this condition be a minimal dietary control group. We emailed participants in this group new lessons weekly for the first 4 weeks and then every 4 weeks thereafter. This group did not get the positive affect regulation and mindful eating materials.

All participants in both groups could contact the first author by phone or email as needed with questions. A coauthor (SK) was on-call by pager for any urgent medical concerns. All regularly planned emails were sent automatically by custom software, which allowed us to ensure that the emails arrived to participants in a timely manner and removed the need for study staff to oversee this process.

Measures

All measures were assessed at baseline before randomization and at 16 and 32 weeks after the intervention began.

Metabolic Measures

We assessed HbA_{1c} as well as fasted low- and high-density lipoprotein cholesterol (LDL and HDL) and triglycerides at a commercial Clinical Laboratory Improvement Amendments-certified laboratory (LabCorp; Laboratory Corporation of America Holdings, Burlington, NC, USA).

Body Weight

Participants recruited at the start of the study had their body weight measured at a US HealthWorks Medical Group (Valencia, CA, USA) location, near wherever they lived. Due to measurement concerns (eg, participants were asked their weight instead of actually being weighed), we then mailed the participants the EatSmart Digital Bathroom Scale. At each critical time point, participants emailed study staff a photo of their feet and digitally displayed weight while they stood on the digital scale.

Psychological Self-Report

Participants completed the Diabetes Distress Scale [30], a measure of upset related to having diabetes. We assessed the subjective experience of each diet by asking, “How much do you like how you feel on your diet?” and “How much do you think your diet has improved your physical health?” all rated from 1 (not at all) to 7 (very much so). We further asked, “How often do you cheat on your diet?” rated from 1 (not at all) to 7 (very often) and “How hard is it to stay on your diet?” rated from 1 (not at all) to 7 (very difficult).

We measured depressive symptoms with a 20-item scale, the Center for Epidemiologic Studies Depression Scale (CESD) [31], with higher scores reflecting greater symptoms over the past week. Following past research, we also separately examined four items that tap into positive affect, including “I felt hopeful about the future” and “I was happy.” Higher scores reflect greater positive affect.

Participants completed the Modified Differential Emotions Scale (mDES) [32], which gauges negative and positive mood. This version of the mDES asked participants to recall the past week and rate how often they had experienced particular emotions, rated from 1 (not at all) to 9 (all the time). The positive emotions subscale consists of amusement, awe, compassion, contentment, gratitude, hope, interest, happy, love, and pride. The negative emotions subscale consists of anger, contempt, disgust, embarrassment, anxiety, guilt, sadness, boredom, and loneliness.

Physical Self-Report

We assessed physical symptoms with an adaptation of the Health Symptom Checklist, a short, face-valid measure of physical symptoms [33], rated from 1 (not at all) to 4 (very often) for how often over the past week they had experienced a variety of physical symptoms. We used a subscale of the Short Form Health Survey [34], a well-validated and extensively used measure of health-related quality of life, to assess vitality (energy and fatigue).

Dietary Self-Report

We assessed dietary composition using the free online application MyFitnessPal [35], which has a vast database of foods and has been or is being used in other clinical trials [36-38]. Even so, its database is partially user-generated and results may be prone to error. Therefore, the dietary self-report results should not be considered validated. At each of the main time points, participants reported on what they had eaten over two weekdays and one weekend day, which we then averaged into one composite measure.

Statistical Analyses

The primary statistical analyses were performed using Stata IC software version 14.1 (StataCorp LP, College Station, TX, USA) setting two-tailed alpha to reject the null hypothesis at .05. Our experimental design randomized participants to one of two groups (intervention: n=12; control: n=13) participating in a 32-week online dietary and lifestyle intervention designed for weight and HbA_{1c} reductions with primary outcomes (HbA_{1c} and body weight) measured at three time points (baseline, week 16, week 32). All our main outcomes were continuously scaled and were appropriately analyzed with parametric statistical techniques, and all statistical assumptions were tested prior to interpreting results. The data met the distributional requirements for the techniques employed without requiring data transformations, model adjustments (eg, random slope terms, heteroscedasticity adjustments), or nonlinear modeling.

Participants’ repeated measures outcomes were submitted to separate mixed-effects linear regression analyses with fixed effect terms comparing baseline (preintervention) to each of the two subsequent observations made postintervention (weeks 16 and 32), the main effect for group, and most importantly, the simple interaction effects comparing the relative change by group at each postintervention assessment, relative to baseline. Random y-intercept terms were included to accommodate for the repeated measures experimental design. Our analysis of total caloric intake, net carbohydrates, fat, and sugar required log transformations prior to analysis to normalize model residuals; out of a total possible 63 observations, we eliminated one triglyceride, two LDL, and one calorie observations that were overly influential outliers.

For all the self-reported ratings of the subjective experience of the diet, we assessed differences between the groups using Cohen *d*. For all results involving ratios, we used a two-tailed Fisher exact test to assess significance. Means and confidence intervals are reported in their original units for all variables.

Results

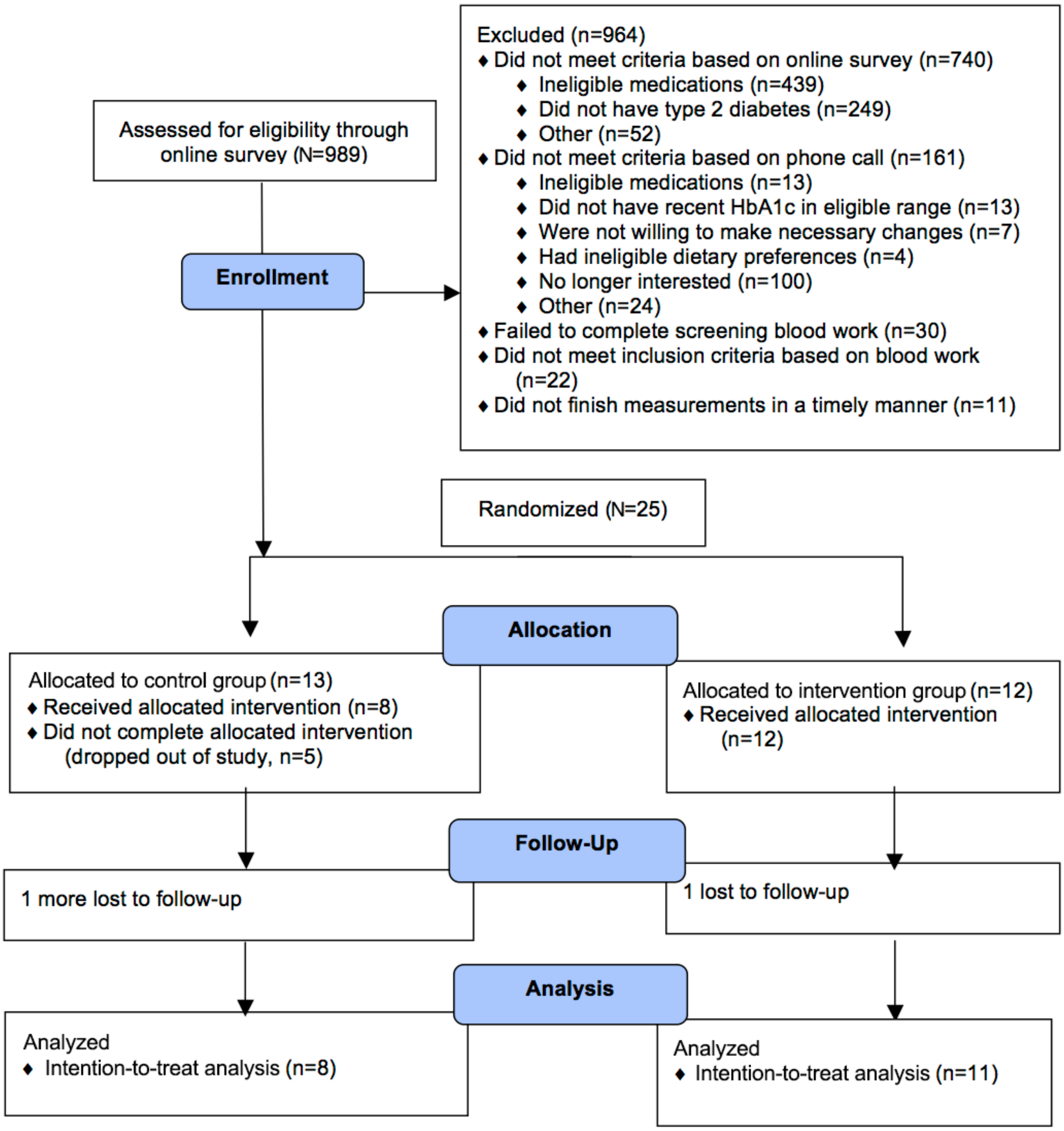
We enrolled and randomized 25 participants to the intervention (n=12) or control (n=13) group (Figure 1). A large number of individuals who took our initial online survey were ineligible because they did not have type 2 diabetes (n=249), were taking diabetes medications other than metformin (n=404), or had definite plans to begin taking insulin (n=35). Randomized participants included men and women, of several different types of ethnic and racial backgrounds (although about half were white), with an average duration of diagnosed type 2 diabetes of approximately 5 years, and a mean baseline HbA_{1c} of approximately 7% (Table 1).

Table 1. Baseline participant characteristics (N=25).

Baseline characteristics	Intervention group (n=12)	Control group (n=13)
Sex, n (%)		
Male	6 (50)	4 (31)
Female	6 (50)	9 (69)
Age (years), mean (SD)	53.0 (10.2)	58.2 (6.7)
Race/Ethnicity, n (%)		
Asian/Pacific Islander	2 (17)	2 (15)
Black	3 (25)	0 (0)
White	7 (58)	8 (62)
Latino/a	2 (17)	5 (38)
Duration of diabetes (years), mean (SD)	5.3 (4.1)	5.7 (3.7)
HbA_{1c} (%), mean (SD)	7.1 (0.4)	7.2 (0.3)
Weight (kg), mean (SD)	109.7 (24.9)	90.9 (16.4)
Triglycerides (mg/dL), mean (SD)	174.1 (79.4)	151.5 (87.1)
HDL cholesterol (mg/dL), mean (SD)	45.7 (15.0)	53.9 (12.7)
LDL cholesterol (mg/dL), mean (SD)	96.9 (30.4)	90.5 (29.0)
Diabetes-related distress, mean (SD)	1.9 (0.8)	2.4 (1.2)
CES-Depression, mean (SD)	10.5 (7.7)	9.8 (7.4)
CES-D Positive Affect, mean (SD)	10.2 (2.3)	10.2 (2.2)
DES Negative Affect, mean (SD)	2.8 (1.3)	2.7 (1.4)
DES Positive Affect, mean (SD)	6.5 (1.1)	6.2 (1.5)
Vitality (SF-36 subscale), mean (SD)	53.3 (16.4)	49.2 (20.1)
Total kilocalories, mean (SD)	1768.5 (626.6)	1749.1 (322.2)
Total grams of nonfiber carbohydrates, mean (SD)	163.6 (86.7)	152.0 (58.9)
Total grams of fat, mean (SD)	77.1 (41.4)	81.3 (27.3)
Total grams of protein, mean (SD)	83.3 (18.0)	74.5 (17.2)
Total grams of sugar, mean (SD)	50.6 (33.8)	44.9 (23.8)

Trial retention differed by group. Dropout was higher in the control group: 16-week dropout for the intervention group was zero of 12 (0%) and 5 of 13 (39%) for the control group (*P*=.04); 32-week dropout for the intervention group was 1 of 12 (8%) and 6 of 13 (46%) for the control group (*P*=.07). One participant in each group reported experiencing an event that they believed was caused by hypoglycemia (one in the control group was likely from eating very few calories and the other in the intervention group was after taking a dose of metformin).

Figure 1. Study participant flowchart for online study.



Metabolic Measures

Glycated Hemoglobin

There were significantly greater reductions in HbA_{1c} for the intervention group relative to the control group at both 16 (*P*=.01) and 32 (*P*=.002) weeks. Reductions in HbA_{1c} were approximately twice as large in the intervention versus the control group (intervention group: estimated marginal mean [EMM] –0.9% at 16 weeks and EMM –0.8% at 32 weeks; control group: EMM –0.5% at 16 weeks and EMM –0.4% at 32 weeks; Table 2,Figure 2). At both 16 and 32 weeks, a greater percentage of participants in the intervention group lowered their HbA_{1c} to less than 6.5%, the cutoff for type 2 diabetes, compared to the percentage of participants in the control group (intervention group: 9/12, 75% at 16 weeks and 6/11, 55% at 32 weeks; control group: 1/8, 13% at 16 weeks and 0/8, 0% at 32 weeks; Table 3). We redid these analyses using participants' baseline body weight as covariates in the model. The results were nearly identical; therefore, we present the simpler unadjusted model.

Table 2. Estimated marginal mean (EMM) changes from baseline to 16 and 32 weeks.[a]

Outcomes	Intervention group, EMM (95% CI)[b]	Control group, EMM (95% CI)[c]	Difference between groups, EMM (95% CI)	*P*
HbA$_{1c}$ (%)				
16 weeks	–0.9 (–1.1, –0.6)	–0.5 (–0.8, –0.2)	–0.4 (–0.7, –0.1)	.01
32 weeks	–0.8 (–1.1, –0.6)	–0.3 (–0.6, 0.0)	–0.5 (–0.8, –0.2)	.002
Weight (kg)				
16 weeks	–8.5 (–11.9, –5.2)	–3.9 (–8.0, 0.2)	–4.6 (–8.8, –0.4)	.03
32 weeks	–12.7 (–16.1, –9.2)	–3.0 (–7.3, 1.3)	–9.6 (–14.0, –5.3)	<.001
Triglycerides (mg/dL)				
16 weeks	–35.5 (–65.7, –5.2)	–17.4 (–55.2, 20.4)	–18.1 (–56.1, 19.9)	.35
32 weeks	–60.1 (–91.3, –28.9)	–6.2 (–46.0, 33.6)	–53.9 (–93.6, –14.2)	.01
HDL cholesterol (mg/dL)				
16 weeks	1.4 (–2.7, 5.6)	–0.3 (–5.3, 4.8)	1.7 (–3.4, 6.8)	.52
32 weeks	4.8 (0.5, 9.1)	0.6 (–4.7, 5.9)	4.1 (–1.2, 9.5)	.13
LDL cholesterol (mg/dL)				
16 weeks	–0.8 (–10.9, 9.4)	1.5 (–11.7, 14.7)	–2.2 (–15.3, 10.8)	.74
32 weeks	–0.3 (–10.8, 10.3)	–6.1 (–19.9, 7.7)	5.9 (–7.8, 19.5)	.40
Diabetes-related distress				
16 weeks	–0.5 (–0.8, –0.1)	–0.3 (–0.7, 0.1)	–0.1 (–0.6, 0.3)	.49
32 weeks	–0.4 (–0.8, 0.0)	–0.4 (–0.8, 0.0)	0.0 (–0.5, 0.5)	.98
CES-Depression				
16 weeks	–3.7 (–7.8, 0.5)	0.8 (–3.7, 5.4)	–4.5 (–9.3, 0.4)	.07
32 weeks	–0.6 (–5.0, 3.7)	–1.0 (–6.0, 4.0)	–0.4 (–4.8, 5.6)	.88
CESD Positive Affect				
16 weeks	8.4 (–5.3, 22.2)	–4.4 (–19.1, 10.3)	12.9 (–2.9, 28.7)	.11
32 weeks	0.5 (–13.6, 14.8)	7.2 (–9.0, 23.4)	–6.6 (–23.5, 10.3)	.45
DES Negative Affect				
16 weeks	–0.7 (1.5, 0.1)	–0.1 (–0.9, 0.8)	–0.6 (–1.5, 0.3)	.19
32 weeks	–0.4 (–1.2, 0.4)	–0.7 (–1.6, 0.2)	0.3 (0.5, –0.6)	.49
DES Positive Affect				
16 weeks	0.5 (–0.3, 1.4)	–0.2 (–1.1, 0.7)	0.7 (–0.3, 1.7)	.15
32 weeks	0.4 (–0.5, 1.2)	0.3 (–0.7, 1.2)	0.1 (–0.9, 1.1)	.82
Vitality (SF-36 subscale)				
16 weeks	13.3 (2.5, 24.02)	2.3 (–9.4, 13.9)	11.0 (–1.4, 23.4)	.08
32 weeks	9.2 (–1.9, 20.4)	11.0 (–1.8, 23.8)	–1.8 (–15.1, 11.6)	.80
Total kilocalories				
16 weeks	–362.9 (–634.7, –91.1)	–300.8 (–594.3, –7.4)	–62.1 (–376.0, 251.7)	.65
32 weeks	–439.3 (–719.4, –159.3)	–216.6 (–559.2, 125.9)	–222.7 (–569.9, 124.5)	.13
Total grams of nonfiber carbohydrates				
16 weeks	–123.2 (–167.2, –79.2)	–27.03 (–75.4, 21.3)	–19.2 (–147.4, –44.9)	<.001
32 weeks	–122.7 (–167.9, –77.5)	–14.8 (–71.0, 41.5)	–107.9 (–164.6, –51.3)	<.001
Total grams of fat				

Outcomes	Intervention group, EMM (95% CI)[b]	Control group, EMM (95% CI)[c]	Difference between groups, EMM (95% CI)	*P*
16 weeks	–7.8 (–36.3, 20.7)	–18.7 (–48.8, 11.4)	10.9 (–21.7, 43.4).	.42
32 weeks	–4.0 (–33.2, 25.3)	–23.7 (–58.5, 11.1)	19.8 (–15.9, 55.5)	.13
Total grams of protein				
16 weeks	–0.5 (–14.0, 13.0)	–0.2 (–15.3, 14.9)	–0.3 (–16.2, 15.6)	.97
32 weeks	–1.6 (–15.5, 12.3)	–0.1 (–17.7, 17.6)	–1.5 (–19.2, 16.1)	.86
Total grams of sugar				
16 weeks	–37.2 (–53.8, –20.5)	–4.4 (–22.9, 14.1)	–32.7 (–52.3, –13.2)	<.001
32 weeks	–32.5 (–49.7, –15.4)	–0.3 (–21.9, 21.3)	–32.2 (–53.9, –10.6)	<.001

[a] Data are estimated marginal means and 95% confidence intervals by linear mixed-effects model analysis.

[b] Total analyzed in intervention group: n=12 for week 16 and n=11 for week 32.

[c] Total analyzed in control group: n=9 for week 16 and n=8 for week 32.

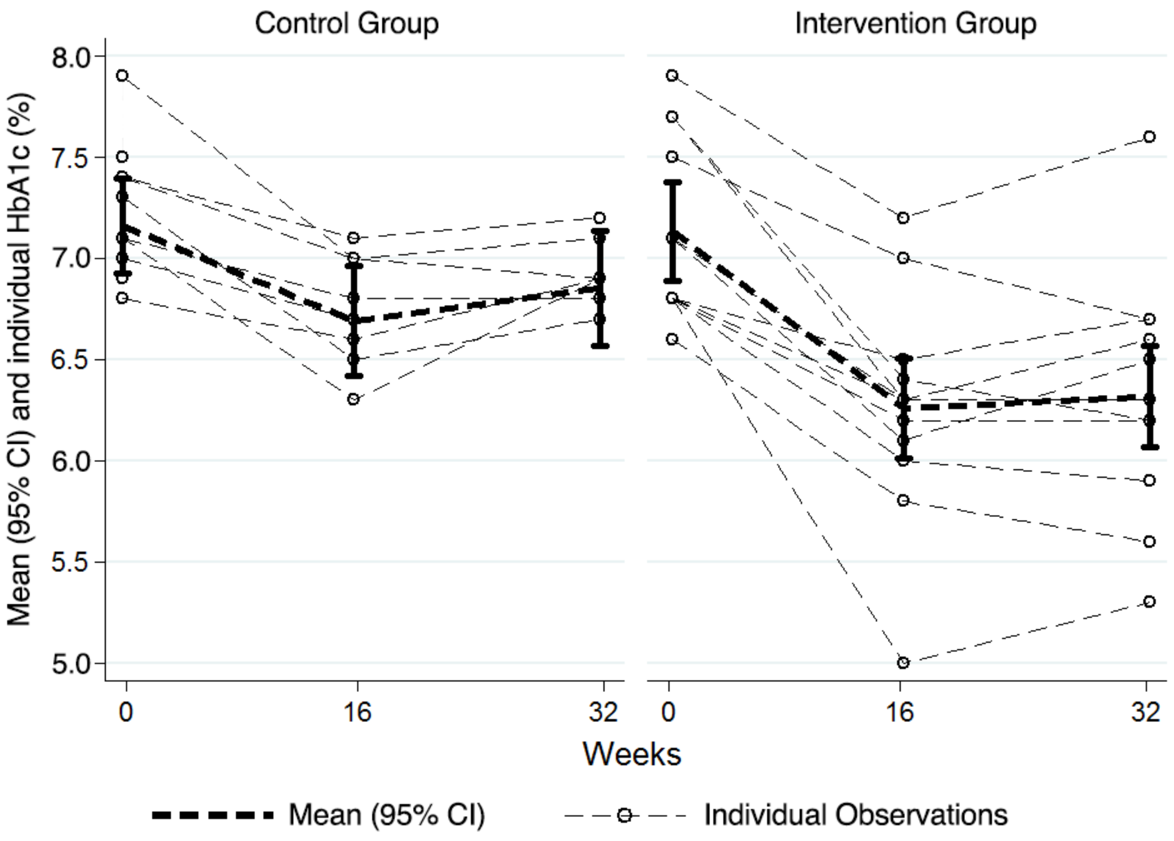
Table 3. Percentage of people meeting HbA_{1c} and weight change thresholds.

HbA_{1c} and weight outcomes	Intervention group[a]	Control group[b]	Difference between groups	*P*
Participants with final HbA_{1c} <6.5%, n (%)				
16 weeks	9 (75%)	1 (13%)		.02
32 weeks	6 (55%)	0 (0%)		.02
Weight (% of initial weight), mean (SD)				
16 weeks	–7.8 (3.6)	–4.2 (3.7)	–3.6 (–7.1, –0.1)	.04
32 weeks	–12.0 (7.3)	–2.5 (4.6)	–9.5 (–16.1, –2.9)	.01
Participants achieving a 5% weight loss, n (%)				
16 weeks	10 (83%)	3 (38%)		.06
32 weeks	10 (90%)	2 (29%)		.01

[a] Total in intervention group: n=12 at 16 weeks and n=11 at 32 weeks.

[b] Total in control group: n=8 at 16 weeks and n=8 at 32 weeks for participants with HbA_{1c} <6.5%, and n=7 at 32 weeks for participants achieving a 5% weight loss.

Figure 2. Mean and individual body weight (in kilograms) for the intervention and control groups at baseline and at 16 and 32 weeks. Bars represent 95% confidence intervals of the mean. Dashed lines reflect individual participants; darker lines represent each group mean.



Body Weight

We also found significantly greater reductions in body weight for participants in the intervention group relative to the control group at weeks 16 (P=.03) and 32 (P<.001). For example, at 32 weeks, participants in the intervention group lost more weight (EMM –12.7 kg, 95% CI –16.1 to –9.2 kg) than participants in the control group (EMM –3.0 kg, 95% CI –7.3 to 1.3 kg; P<.001) (Table 2,Figure 3). At both 16 and 32 weeks, more than double the percentage of participants in the intervention group lost at least 5% of their body weight compared participants in the control group (intervention group: 10/12, 83.3% at 16 weeks and 10/11, 90.1% at 32 weeks; control group: 3/8, 37.5% at 16 weeks and 2/8, 28.6% at 32 weeks; Table 3).

We examined the intersection of weight and HbA_{1c} changes over time for each participant (Figure 4). Participants in the intervention group tended to show strong downward and leftward trajectories, especially from baseline to 16 weeks, reflecting a strong initial loss in weight and HbA_{1c}, whereas those trajectories reflected less simultaneous change for most participants in the control group.

Figure 3. Mean and individual body weight (in kilograms) for the intervention and control groups at baseline and at 16 and 32 weeks. Bars represent 95% confidence intervals of the mean. Dashed lines reflect individual participants; darker lines represent each group mean.

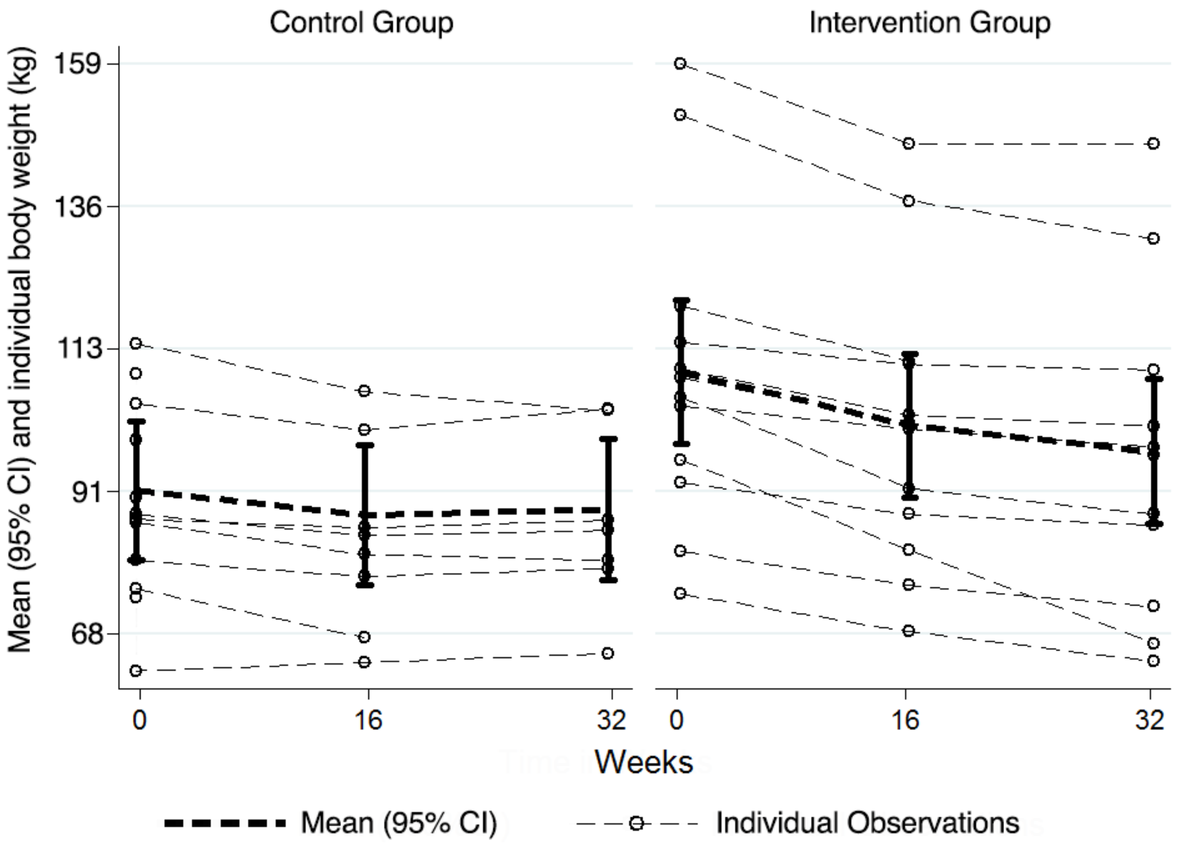
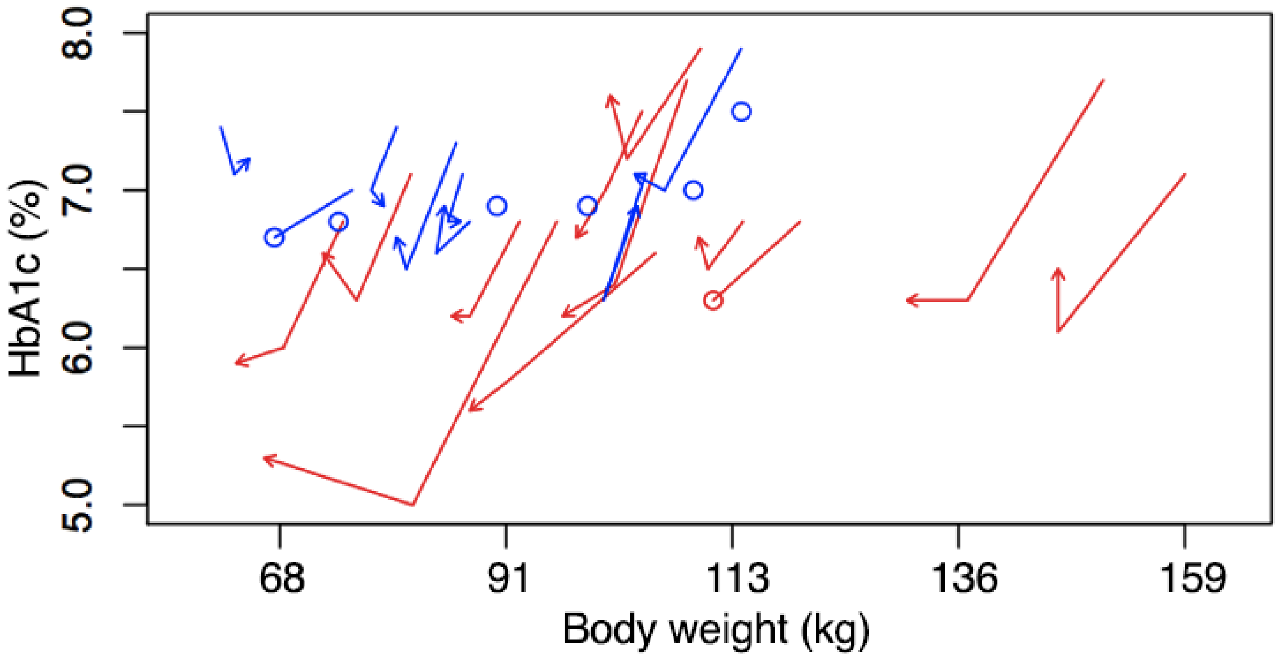


Figure 4. Body weight and HbA_{1c} plotted for each participant separately for each of the three time periods (0, 16, and 32 weeks). Red lines represent the intervention participants; blue lines represent the control participants. Lines that end in an O reflect dropouts (and missing data). Lines that end in an arrow show participants who completed the study.



Cholesterol and Triglycerides

The intervention was also more effective at reducing triglycerides from baseline relative to the control; however, the effect was significant only at the 32-week time point (*P*=.01). Both HDL and LDL data revealed no effects between groups or differences from baseline within each group.

Psychological Self-Report

Diabetes-Related Distress

We found no statistically significant effects on this measure in the intervention group relative to the control group (Table 2).

Depressive Symptoms, Affect, and Vitality

We found no statistically significant effects on these measures in the intervention group relative to the control group (Table 2).

Subjective Experience of the Diets

Compared to the control group, participants in the intervention group rated themselves as less likely to cheat on their assigned diet at 16 and 32 weeks, with a large effect size of at least a Cohen *d*=–1.0 (Table 4). Participants in the intervention group also rated their diet less difficult to stick to, better liked how they felt on the diet, and were more likely to think that their diet improved their physical health, all with medium to large effect sizes.

Table 4. Self-reported ratings of subjective experience of the diets.

Self-reported ratings	Intervention group, mean (SD)	Control group, mean (SD)	Cohen *d* between groups
Overall self-rating of how much they like how they feel on the diet			
16 weeks	5.9 (1.1)	5.2 (1.3)	0.6
32 weeks	6.2 (1.0)	4.9 (2.3)	0.8
Overall self-rating of how much they think the diet improved their physical health			
16 weeks	6.2 (1.0)	5.3 (1.8)	0.6
32 weeks	6.2 (0.9)	5.1 (2.5)	0.5
Overall self-rating of likelihood of cheating on diet			
16 weeks	2.7 (1.4)	3.9 (0.9)	–1.0
32 weeks	3.4 (1.1)	5.0 (0.8)	–1.7
Overall self-rating of difficulty of staying on diet			
16 weeks	2.7 (1.5)	4.0 (1.3)	–0.9
32 weeks	3.2 (1.5)	4.0 (1.7)	–0.5

Physical Self-Report

Compared to participants in the control group, participants in the intervention group reported greater reductions in headache symptoms, bloating, and gas at week 16, as well as greater increases in constipation symptoms at week 16 (all with a large Cohen *d* effect size of at least 0.9 between groups; Multimedia Appendices 1 and 2.

Dietary Self-Reports

At both 16 and 32 weeks, compared to the control group, the intervention group reported eating fewer grams of nonfiber carbohydrates and grams of sugar. Grams of fat and protein did not show any group effects (Table 2).

The dietary measurements suggest that participants in both groups were, on average, adherent to their assigned diet. Participants in the intervention group ate the recommended daily grams of nonfiber carbohydrates and participants in the control group ate an expected percentage of their total calories from carbohydrates. In the intervention group, daily grams of nonfiber carbohydrates lowered from a mean of 163.6 (SD 86.7) grams at baseline to a mean of 40.4 (SD 45.9) grams at 16 weeks and a mean of 43.5 (SD 33.9) grams at 32 weeks, suggesting that participants were, on average, adherent to their assigned intervention diet. Although participants in the control group were not asked to reach a particular daily target for grams of nonfiber carbohydrates, their daily intake lowered from a mean of 152.0 (SD 58.9) grams at baseline to a mean of 127.1 (SD 40.2) grams at 16 weeks and a mean of 144.8 (SD 33.7) grams at 32 weeks.

In the intervention group, the mean percentage of calories from total carbohydrates changed from a baseline mean of 39.6% (SD 10.4%) to a mean of 15.5% (SD 13.0%) at 16 weeks and a mean of 18.5% (SD 12.8%) at 32 weeks. In the control group, the mean percentage of calories from total carbohydrates changed from a baseline mean of 37.6% (SD 10.3%) to a mean of 40.9% (SD 6.3%) at 16 weeks and a mean of 43.0% (SD 9.1%) at 32 weeks. This percentage of calories from carbohydrates in the both groups suggest that participants were, on average, adherent their assigned control diet; the overall target of percentage of calories from carbohydrates was expected to be approximately less than 20% in the intervention group and about 50% in the control group.

Medication Changes

The changes in metformin dosages were similar between groups. At 32 weeks, in the intervention group, metformin medications was decreased in one participant, increased in two participants, and unchanged in eight participants. In the control group, metformin dosage was decreased in two participants, increased in one participant, and unchanged in four participants. We had limited room to see differences in medication changes because we only enrolled participants on metformin, a drug that is safe enough and has a low enough risk of hypoglycemia that physicians do not quickly change its dosage.

Discussion

Principal Results

Our results show that participants randomly assigned to the very low-carbohydrate ketogenic diet and lifestyle recommendations (intervention) group had a variety of health benefits including lower HbA_{1c}, body weight, and triglyceride levels, compared to those assigned to the control group (the plate method diet).

Our results are similar to those from our previous in-person trial of these very low-carbohydrate ketogenic diet and lifestyle recommendations. This online study differed from the in-person one due to its recruitment approach (national for this online study; in San Francisco, CA, for the in-person study), and allowable diabetes medications (none or just metformin for this online study; none, metformin, and/or sulfonylureas or dipeptidyl peptidase-4 inhibitor for the in-person study). These differences suggest that the online program might be applicable to overweight individuals with type 2 diabetes living across the United States, not currently taking multiple medications for their diabetes, who are motivated to make dietary changes, and have access to the Internet.

Limitations

The ability to generalize from the results of this study is limited by its size, targeted population, and length of follow-up time. We had to screen a large number of participants in order to find eligible participants. Of those screened, 26% filled out the online survey but did not have type 2 diabetes (according to their own self-report), 47% were taking or planning to take medications that made them ineligible, 10% were not interested once they heard more about the study from the study staff online, and only a few (<1%) reported having dietary preferences counter to those on the possible assigned diets. A larger trial with a longer follow-up is needed to better understand the durability of the effects on glycemic control and weight in a broader population.

By the end of the trial, we retained 92% of participants in the intervention group, compared to 54% of the control group. This difference could have been because the intervention group's program had more sessions and included behavioral adherence strategies, which may have made their program more engaging. In addition, glycemic control and weight loss were lower in the intervention group; some participants in the control group expressed frustration that their glycemic control or weight loss was not as much as they would like, and thus they decided to not continue with the control program. Participants in the intervention group rated themselves as less likely to cheat on their assigned diet, compared to participants in the control group. Perhaps this difference in likelihood to cheat also suggests that the intervention program was easier to adhere to (possibly due to the diet or possibly due to the extra supports included in the intervention program).

Comparison With Prior Work

An innovative aspect of this program was remote monitoring of glycemic control, body weight, and other outcomes, suggesting that although these online program participants never met the researchers or study staff in-person, we were still able to measure and improve outcomes. Several other online interventions have successfully improved glycemic control and reduced body weight in adults, although their ability to retain participants was mixed. For example, one online program for individuals with prediabetes was based on the Diabetes Prevention Program, and it likely taught participants a lower-calorie, lower-fat diet, although this is not explicitly mentioned in the publication. After 12 months, participants' HbA_{1c} was reduced by 0.4% and they had lost 4.8% of body weight. However, only 45% of 220 participants had follow-up HbA_{1c} values [39]. In a completely online program for overweight individuals that taught participants to follow a lower-calorie and lower-fat diet, participants in the active intervention group lost a mean 5.6% of their body weight 6 months after baseline. Of 77 participants assigned to the active intervention group, 70% were retained [40].

Not all online trials are effective. In an online self-management program for individuals with type 2 diabetes that recommended “healthy eating,” participants in the active intervention groups (with or without extra follow-up calls and visits) had not significantly reduced their HbA_{1c} levels or lost weight at 12 months after baseline. Of 331 participants assigned to either intervention group, 72% were retained [41]. In an online self-care intervention for individuals with type 2 diabetes that also recommended “healthy eating,” participants assigned to the active intervention group did not show changes in their HbA_{1c} at 6 months after baseline. Of 491 participants who began the program, 80% had 6-month outcome data [42]. Thus, the retention rate of our active intervention group (92%) was good compared to other previous online trials, as was our ability to bring about changes in glycemic control and weight in that group.

In previous trials of very low-carbohydrate diet programs in adults with type 2 diabetes that were at least 3 months or longer [7,38,43-58], researchers followed participants for a mean of 12 months. All interventions were in-person. On average, HbA_{1c} dropped 1.0% (median –0.8%). Both mean and median body weight lost was –8%. Thus, the results from our online very low-carbohydrate intervention program replicate or improve on past results, given the fact that HbA_{1c} dropped by 0.8% and body weight reduced by 12.0%.

Conclusions

Our results lend continued support for the idea that our program's recommendations to follow a very low-carbohydrate ketogenic diet and make lifestyle changes is promising and can bring about improved health outcomes in overweight individuals with type 2 diabetes. Future work should examine how robust these results are with larger, more diverse participants; determine whether more robust psychological or other intervention support could improve dietary adherence; track whether the positive health effects are sustained over time; and, through more thorough implementation research, whether and how such an online intervention can dovetail with existing in-person health care teams. The online delivery of this approach gives it the potential to have wider impact in the treatment of type 2 diabetes.

Acknowledgments

Thanks very much to our dedicated participants as well as Patricia Moran and Michael Cohn for their help. The research was supported by a grant from the Mount Zion Health Fund. Laura Saslow and Ashley Mason were supported by National Institutes of Health (NIH) grant T32AT003997 from the National Center for Complementary and Integrative Health (NCCIH). Laura Saslow was also supported by funding from the William K Bowes, Jr Foundation and the NIH (K01 from the National Institute of Diabetes and Digestive and Kidney Diseases, DK107456). Ashley Mason was also supported by the NIH (K23 from the National Heart, Lung, and Blood Institute, HL133442). Judith Moskowitz was supported by NIH grant K24 MH093225 from the National Institute of Mental Health. Frederick Hecht was supported by NIH grant K24 AT007827 from NCCIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

Frederick Hecht is on the Scientific Advisory Board for Virta Health. No other author declares any conflict of interest.

Multimedia Appendix 1

Baseline values of physical symptoms.

[PDF File (Adobe PDF File), 7KB - jmir_v19i2e36_app1.pdf]

Multimedia Appendix 2

Change in symptoms from baseline.

[PDF File (Adobe PDF File), 10KB - jmir_v19i2e36_app2.pdf]

References

1. American DA. Economic costs of diabetes in the US in 2012. Diabetes Care 2013 Apr;36(4):1033-1046 [FREE Full text] [doi: 10.2337/dc12-2625] [Medline: 23468086]
2. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. Lancet 2011 Jul 2;378(9785):31-40. [doi: 10.1016/S0140-6736(11)60679-X] [Medline: 21705069]
3. Feinman RD, Pogozelski WK, Astrup A, Bernstein RK, Fine EJ, Westman EC, et al. Dietary carbohydrate restriction as the first approach in diabetes management: critical review and evidence base. Nutrition 2015 Jan;31(1):1-13 [FREE Full text] [doi: 10.1016/j.nut.2014.06.011] [Medline: 25287761]
4. Volek JS, Phinney SD, Forsythe CE, Quann EE, Wood RJ, Puglisi MJ, et al. Carbohydrate restriction has a more favorable impact on the metabolic syndrome than a low fat diet. Lipids 2009 Apr;44(4):297-309. [doi: 10.1007/s11745-008-3274-2] [Medline: 19082851]
5. Ruskin DN, Masino SA. The nervous system and metabolic dysregulation: emerging evidence converges on ketogenic diet therapy. Front Neurosci 2012;6:33 [FREE Full text] [doi: 10.3389/fnins.2012.00033] [Medline: 22470316]
6. Bistrian BR, Blackburn GL, Flatt JP, Sizer J, Scrimshaw NS, Sherman M. Nitrogen metabolism and insulin requirements in obese diabetic adults on a protein-sparing modified fast. Diabetes 1976 Jun;25(6):494-504. [Medline: 1278601]
7. Dashti HM, Mathew TC, Khadada M, Al-Mousawi M, Talib H, Asfar SK, et al. Beneficial effects of ketogenic diet in obese diabetic subjects. Mol Cell Biochem 2007 Aug;302(1-2):249-256. [doi: 10.1007/s11010-007-9448-z] [Medline: 17447017]
8. Kirk JK, Graves DE, Craven TE, Lipkin EW, Austin M, Margolis KL. Restricted-carbohydrate diets in patients with type 2 diabetes: a meta-analysis. J Am Diet Assoc 2008 Jan;108(1):91-100. [doi: 10.1016/j.jada.2007.10.003] [Medline: 18155993]

9. Nielsen JV, Joensson E. Low-carbohydrate diet in type 2 diabetes. Stable improvement of bodyweight and glycemic control during 22 months follow-up. Nutr Metab (Lond) 2006;3:22 [FREE Full text] [doi: 10.1186/1743-7075-3-22] [Medline: 16774674]
10. Accurso A, Bernstein RK, Dahlqvist A, Draznin B, Feinman RD, Fine EJ, et al. Dietary carbohydrate restriction in type 2 diabetes mellitus and metabolic syndrome: time for a critical appraisal. Nutr Metab (Lond) 2008;5:9 [FREE Full text] [doi: 10.1186/1743-7075-5-9] [Medline: 18397522]
11. Santos FL, Esteves SS, da Costa PA, Yancy WS, Nunes JP. Systematic review and meta-analysis of clinical trials of the effects of low carbohydrate diets on cardiovascular risk factors. Obes Rev 2012 Nov;13(11):1048-1066. [doi: 10.1111/j.1467-789X.2012.01021.x] [Medline: 22905670]
12. Markwald RR, Melanson EL, Smith MR, Higgins J, Perreault L, Eckel RH, et al. Impact of insufficient sleep on total daily energy expenditure, food intake, and weight gain. Proc Natl Acad Sci U S A 2013 Apr 2;110(14):5695-5700 [FREE Full text] [doi: 10.1073/pnas.1216951110] [Medline: 23479616]
13. Broussard JL, Ehrmann DA, Van CE, Tasali E, Brady MJ. Impaired insulin signaling in human adipocytes after experimental sleep restriction: a randomized, crossover study. Ann Intern Med 2012 Oct 16;157(8):549-557 [FREE Full text] [doi: 10.7326/0003-4819-157-8-201210160-00005] [Medline: 23070488]
14. Spiegel K, Knutson K, Leproult R, Tasali E, Van Cauter E. Sleep loss: a novel risk factor for insulin resistance and Type 2 diabetes. J Appl Physiol (1985) 2005 Nov;99(5):2008-2019 [FREE Full text] [doi: 10.1152/japplphysiol.00660.2005] [Medline: 16227462]
15. Spiegel K, Tasali E, Penev P, Van CE. Brief communication: sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. Ann Intern Med 2004 Dec 7;141(11):846-850. [Medline: 15583226]
16. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
17. Saslow L, Cohn M, Moskowitz J. Positive affect interventions to reduce stress: harnessing benefit while avoiding Pollyanna. In: Gruber J, Moskowitz J, editors. The Light and Dark Sides of Positive Emotion. New York: Oxford University Press; Jan 23, 2014:515-532.
18. Daubenmier J, Kristeller J, Hecht FM, Maninger N, Kuwata M, Jhaveri K, et al. Mindfulness intervention for stress eating to reduce cortisol and abdominal fat among overweight and obese women: an exploratory randomized controlled study. J Obes 2011;2011:651936 [FREE Full text] [doi: 10.1155/2011/651936] [Medline: 21977314]
19. Daubenmier J, Moran PJ, Kristeller J, Acree M, Bacchetti P, Kemeny ME, et al. Effects of a mindfulness-based weight loss intervention in adults with obesity: a randomized clinical trial. Obesity (Silver Spring) 2016 Apr;24(4):794-804 [FREE Full text] [doi: 10.1002/oby.21396] [Medline: 26955895]
20. Saslow LR, Kim S, Daubenmier JJ, Moskowitz JT, Phinney SD, Goldman V, et al. A randomized pilot trial of a moderate carbohydrate diet compared to a very low carbohydrate diet in overweight or obese individuals with type 2 diabetes mellitus or prediabetes. PLoS One 2014;9(4):e91027 [FREE Full text] [doi: 10.1371/journal.pone.0091027] [Medline: 24717684]
21. LaBrie JW, Quinlan T, Schiffman JE, Earleywine ME. Performance of alcohol and safer sex change rulers compared with readiness to change questionnaires. Psychol Addict Behav 2005 Mar;19(1):112-115. [doi: 10.1037/0893-164X.19.1.112] [Medline: 15783287]
22. Bogg T, Roberts BW. Conscientiousness and health-related behaviors: a meta-analysis of the leading behavioral contributors to mortality. Psychol Bull 2004 Nov;130(6):887-919. [doi: 10.1037/0033-2909.130.6.887] [Medline: 15535742]
23. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. J Res Pers 2003 Dec;37(6):504-528. [doi: 10.1016/S0092-6566(03)00046-1]
24. Gearhardt AN, Corbin WR, Brownell KD. Preliminary validation of the Yale Food Addiction Scale. Appetite 2009 Apr;52(2):430-436. [doi: 10.1016/j.appet.2008.12.003] [Medline: 19121351]
25. Stice E, Telch CF, Rizvi SL. Development and validation of the Eating Disorder Diagnostic Scale: a brief self-report measure of anorexia, bulimia, and binge-eating disorder. Psychol Assess 2000 Jun;12(2):123-131. [Medline: 10887758]
26. Garber AJ, Menzel PH, Boden G, Owen OE. Hepatic ketogenesis and gluconeogenesis in humans. J Clin Invest 1974 Oct;54(4):981-989 [FREE Full text] [doi: 10.1172/JCI107839] [Medline: 4430728]
27. Forsythe CE, Phinney SD, Fernandez ML, Quann EE, Wood RJ, Bibus DM, et al. Comparison of low fat and low carbohydrate diets on circulating fatty acid composition and markers of inflammation. Lipids 2008 Jan;43(1):65-77. [doi: 10.1007/s11745-007-3132-7] [Medline: 18046594]
28. Kristeller JL, Wolever RQ. Mindfulness-based eating awareness training for treating binge eating disorder: the conceptual foundation. Eat Disord 2011;19(1):49-61. [doi: 10.1080/10640266.2011.533605] [Medline: 21181579]
29. Wolff K, Cavanaugh K, Malone R, Hawk V, Gregory BP, Davis D, et al. The Diabetes Literacy and Numeracy Education Toolkit (DLNET): materials to facilitate diabetes education and management in patients with low literacy and numeracy skills. Diabetes Educ 2009;35(2):233-238 [FREE Full text] [doi: 10.1177/0145721709331945] [Medline: 19240246]
30. Polonsky WH, Fisher L, Earles J, Dudl RJ, Lees J, Mullan J, et al. Assessing psychosocial distress in diabetes: development of the diabetes distress scale. Diabetes Care 2005 Mar;28(3):626-631. [Medline: 15735199]

31. Roberts RE, Vernon SW. The Center for Epidemiologic Studies Depression Scale: its use in a community sample. Am J Psychiatry 1983 Jan;140(1):41-46. [doi: 10.1176/ajp.140.1.41] [Medline: 6847983]
32. Fredrickson BL, Tugade MM, Waugh CE, Larkin GR. What good are positive emotions in crises? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. J Pers Soc Psychol 2003 Feb;84(2):365-376 [FREE Full text] [Medline: 12585810]
33. Bartone PT, Ursano RJ, Wright KM, Ingraham LH. The impact of a military air disaster on the health of assistance workers. A prospective study. J Nerv Ment Dis 1989 Jun;177(6):317-328. [Medline: 2723619]
34. 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]
35. Yusof A, Iahad N. Review on online and mobile weight loss management system for overcoming obesity. 2012 Presented at: International Conference on Computer & Information Science (ICCIS); June 12-14, 2012; Kuala Lumpur, Malaysia p. 198-203 URL: https://www.researchgate.net/publication/239524687_Review_on_Online_and_Mobile_Weight_Loss_Management_System_for_Overcoming_Obesity [doi: 10.1109/ICCISci.2012.6297239]
36. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-S12. [doi: 10.7326/M13-3005] [Medline: 25402403]
37. Taylor RW, Roy M, Jospe MR, Osborne HR, Meredith-Jones KJ, Williams SM, et al. Determining how best to support overweight adults to adhere to lifestyle change: protocol for the SWIFT study. BMC Public Health 2015;15:861 [FREE Full text] [doi: 10.1186/s12889-015-2205-4] [Medline: 26341820]
38. Willis EA, Szabo-Reed AN, Ptomey LT, Steger FL, Honas JJ, Al-Hihi EM, et al. Distance learning strategies for weight management utilizing social media: A comparison of phone conference call versus social media platform. Rationale and design for a randomized study. Contemp Clin Trials 2016 Feb 12;47:282-288. [doi: 10.1016/j.cct.2016.02.005] [Medline: 26883282]
39. Sepah SC, Jiang L, Peters AL. Translating the diabetes prevention program into an online social network: validation against CDC standards. Diabetes Educ 2014 Apr 10;40(4):435-443. [doi: 10.1177/0145721714531339] [Medline: 24723130]
40. Thomas JG, Leahey TM, Wing RR. An automated internet behavioral weight-loss program by physician referral: a randomized controlled trial. Diabetes Care 2015 Jan;38(1):9-15. [doi: 10.2337/dc14-1474] [Medline: 25404659]
41. Glasgow RE, Kurz D, King D, Dickman JM, Faber AJ, Halterman E, et al. Twelve-month outcomes of an Internet-based diabetes self-management support program. Patient Educ Couns 2012 Apr;87(1):81-92 [FREE Full text] [doi: 10.1016/j.pec.2011.07.024] [Medline: 21924576]
42. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VB, 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]
43. Mayer SB, Jeffreys AS, Olsen MK, McDuffie JR, Feinglos MN, Yancy WS. Two diets with different haemoglobin A1c and antiglycaemic medication effects despite similar weight loss in type 2 diabetes. Diabetes Obes Metab 2014 Jan;16(1):90-93 [FREE Full text] [doi: 10.1111/dom.12191] [Medline: 23911112]
44. Stern L, Iqbal N, Seshadri P, Chicano KL, Daily DA, McGrory J, et al. The effects of low-carbohydrate versus conventional weight loss diets in severely obese adults: one-year follow-up of a randomized trial. Ann Intern Med 2004 May 18;140(10):778-785. [Medline: 15148064]
45. Samaha FF, Iqbal N, Seshadri P, Chicano KL, Daily DA, McGrory J, et al. A low-carbohydrate as compared with a low-fat diet in severe obesity. N Engl J Med 2003 May 22;348(21):2074-2081. [doi: 10.1056/NEJMoa022637] [Medline: 12761364]
46. Hussain TA, Mathew TC, Dashti AA, Asfar S, Al-Zaid N, Dashti HM. Effect of low-calorie versus low-carbohydrate ketogenic diet in type 2 diabetes. Nutrition 2012 Oct;28(10):1016-1021. [doi: 10.1016/j.nut.2012.01.016] [Medline: 22673594]
47. Yancy WS, Foy M, Chalecki AM, Vernon MC, Westman EC. A low-carbohydrate, ketogenic diet to treat type 2 diabetes. Nutr Metab (Lond) 2005 Dec 01;2:34 [FREE Full text] [doi: 10.1186/1743-7075-2-34] [Medline: 16318637]
48. Westman EC, Yancy WS, Mavropoulos JC, Marquart M, McDuffie JR. The effect of a low-carbohydrate, ketogenic diet versus a low-glycemic index diet on glycemic control in type 2 diabetes mellitus. Nutr Metab (Lond) 2008;5:36 [FREE Full text] [doi: 10.1186/1743-7075-5-36] [Medline: 19099589]
49. Shai I, Schwarzfuchs D, Henkin Y, Shahar DR, Witkow S, Greenberg I, Dietary Intervention Randomized Controlled Trial (DIRECT) Group. Weight loss with a low-carbohydrate, Mediterranean, or low-fat diet. N Engl J Med 2008 Jul 17;359(3):229-241. [doi: 10.1056/NEJMoa0708681] [Medline: 18635428]
50. Daly ME, Paisey R, Paisey R, Millward BA, Eccles C, Williams K, et al. Short-term effects of severe dietary carbohydrate-restriction advice in Type 2 diabetes--a randomized controlled trial. Diabet Med 2006 Jan;23(1):15-20. [doi: 10.1111/j.1464-5491.2005.01760.x] [Medline: 16409560]
51. Tay J, Luscombe-Marsh ND, Thompson CH, Noakes M, Buckley JD, Wittert GA, et al. Comparison of low- and high-carbohydrate diets for type 2 diabetes management: a randomized trial. Am J Clin Nutr 2015 Oct;102(4):780-790. [doi: 10.3945/ajcn.115.112581] [Medline: 26224300]

52. Yancy WS, Vernon MC, Westman EC. A pilot trial of a low-carbohydrate, ketogenic diet in patients with type 2 diabetes. Metab Syndr Relat Disord 2003 Sep;1(3):239-243. [doi: 10.1089/154041903322716723] [Medline: 18370668]
53. Davis NJ, Tomuta N, Schechter C, Isasi CR, Segal-Isaacson CJ, Stein D, et al. Comparative study of the effects of a 1-year dietary intervention of a low-carbohydrate diet versus a low-fat diet on weight and glycemic control in type 2 diabetes. Diabetes Care 2009 Jul;32(7):1147-1152 [FREE Full text] [doi: 10.2337/dc08-2108] [Medline: 19366978]
54. Dyson PA, Beatty S, Matthews DR. A low-carbohydrate diet is more effective in reducing body weight than healthy eating in both diabetic and non-diabetic subjects. Diabet Med 2007 Dec;24(12):1430-1435. [doi: 10.1111/j.1464-5491.2007.02290.x] [Medline: 17971178]
55. Vernon MC, Mavropoulos J, Transue M, Yancy WS, Westman EC. Clinical experience of a carbohydrate-restricted diet: effect on diabetes mellitus. Metab Syndr Relat Disord 2003 Sep;1(3):233-237. [doi: 10.1089/154041903322716714] [Medline: 18370667]
56. Iqbal N, Vetter ML, Moore RH, Chittams JL, Dalton-Bakes CV, Dowd M, et al. Effects of a low-intensity intervention that prescribed a low-carbohydrate vs a low-fat diet in obese, diabetic participants. Obesity (Silver Spring) 2010 Sep;18(9):1733-1738 [FREE Full text] [doi: 10.1038/oby.2009.460] [Medline: 20019677]
57. Nielsen JV, Joensson EA. Low-carbohydrate diet in type 2 diabetes: stable improvement of bodyweight and glycemic control during 44 months follow-up. Nutr Metab (Lond) 2008;5:14 [FREE Full text] [doi: 10.1186/1743-7075-5-14] [Medline: 18495047]
58. Guldbrand H, Dizdar B, Bunjaku B, Lindström T, Bachrach-Lindström M, Fredrikson M, et al. In type 2 diabetes, randomisation to advice to follow a low-carbohydrate diet transiently improves glycaemic control compared with advice to follow a low-fat diet producing a similar weight loss. Diabetologia 2012 Aug;55(8):2118-2127 [FREE Full text] [doi: 10.1007/s00125-012-2567-4] [Medline: 22562179]

Abbreviations

CESD: Center for Epidemiologic Studies Depression Scale
HbA_{1c}: glycated hemoglobin
HDL: high-density lipoprotein
LDL: low-density lipoprotein
mDES: Modified Differential Emotions Scale

Edited by G Eysenbach; submitted 27.03.16; peer-reviewed by M Jospe, Y Probst, R Feinman, G Allen; comments to author 06.07.16; revised version received 01.12.16; accepted 05.01.17; published 13.02.17.

Please cite as:
Saslow LR, Mason AE, Kim S, Goldman V, Ploutz-Snyder R, Bayandorian H, Daubenmier J, Hecht FM, Moskowitz JT
An Online Intervention Comparing a Very Low-Carbohydrate Ketogenic Diet and Lifestyle Recommendations Versus a Plate Method Diet in Overweight Individuals With Type 2 Diabetes: A Randomized Controlled Trial
J Med Internet Res 2017;19(2):e36
URL: http://www.jmir.org/2017/2/e36/
doi:10.2196/jmir.5806
PMID:28193599

Original Paper

Barriers to Remote Health Interventions for Type 2 Diabetes: A Systematic Review and Proposed Classification Scheme

Michelle M Alvarado[1], MEng, PhD; Hye-Chung Kum[1,2], MS, MSW, PhD; Karla Gonzalez Coronado[1], BS; Margaret J Foster[3], MS, MPH; Pearl Ortega[1], MEng; Mark A Lawley[1], PE, PhD

[1]Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, United States
[2]Department of Health Policy and Management, School of Public Health, Texas A&M Health Science Center, College Station, TX, United States
[3]Medical Sciences Library, Texas A&M University, College Station, TX, United States

Corresponding Author:
Michelle M Alvarado, MEng, PhD
Department of Industrial and Systems Engineering
Texas A&M University
3131 TAMU
College Station, TX, 77843
United States
Phone: 1 979 458 2350
Fax: 1 979 458 4299
Email: michelle.alvarado@tamu.edu

Abstract

Background: Diabetes self-management involves adherence to healthy daily habits typically involving blood glucose monitoring, medication, exercise, and diet. To support self-management, some providers have begun testing remote interventions for monitoring and assisting patients between clinic visits. Although some studies have shown success, there are barriers to widespread adoption.

Objective: The objective of our study was to identify and classify barriers to adoption of remote health for management of type 2 diabetes.

Methods: The following 6 electronic databases were searched for articles published from 2010 to 2015: MEDLINE (Ovid), Embase (Ovid), CINAHL, Cochrane Central, Northern Light Life Sciences Conference Abstracts, and Scopus (Elsevier). The search identified studies involving remote technologies for type 2 diabetes self-management. Reviewers worked in teams of 2 to review and extract data from identified papers. Information collected included study characteristics, outcomes, dropout rates, technologies used, and barriers identified.

Results: A total of 53 publications on 41 studies met the specified criteria. Lack of data accuracy due to input bias (32%, 13/41), limitations on scalability (24%, 10/41), and technology illiteracy (24%, 10/41) were the most commonly cited barriers. Technology illiteracy was most prominent in low-income populations, whereas limitations on scalability were more prominent in mid-income populations. Barriers identified were applied to a conceptual model of successful remote health, which includes patient engagement, patient technology accessibility, quality of care, system technology cost, and provider productivity. In total, 40.5% (60/148) of identified barrier instances impeded patient engagement, which is manifest in the large dropout rates cited (up to 57%).

Conclusions: The barriers identified represent major challenges in the design of remote health interventions for diabetes. Breakthrough technologies and systems are needed to alleviate the barriers identified so far, particularly those associated with patient engagement. Monitoring devices that provide objective and reliable data streams on medication, exercise, diet, and glucose monitoring will be essential for widespread effectiveness. Additional work is needed to understand root causes of high dropout rates, and new interventions are needed to identify and assist those at the greatest risk of dropout. Finally, future studies must quantify costs and benefits to determine financial sustainability.

(J Med Internet Res 2017;19(2):e28) doi:10.2196/jmir.6382

KEYWORDS

diabetes mellitus, type 2; early medical intervention; biomedical technology; remote sensing technology; terminology as topic

Introduction

Background

Management of type 2 diabetes requires healthy lifestyle habits including diet, medication adherence, and exercise. Thus, patients must practice strong *self-management*, the act of taking responsibility for one's own behavior and well-being. Conventional outpatient therapies fail to address the daily decision-making challenges faced by patients with diabetes [1]. Thus, providers have begun experimenting with remote health to help patients further manage their conditions. *Remote health* is a type of ambulatory health care that allows patients to use technology to collect data and communicate with their health care provider in a different location. We use the term *remote health intervention* to refer to the specific remote health actions and technology employed to improve patient health. This paper focuses on remote health interventions for type 2 diabetes self-management. These remote health interventions can provide a window into the patient's daily activity levels, medication adherence, diet habits, and health vitals. In theory, this transparency enables proactive intervention for poor compliance and emerging risks, assuring better daily health and helping patients avoid hospital visits. This paper assumes that remote health encompasses any type of health care delivered remotely, including telemonitoring, telemedicine, telehealth, eHealth, and mHealth.

Research shows several benefits including high levels of patient satisfaction, positive behavioral changes, and better health outcomes (both physical and mental) [2-4]. Unfortunately, researchers have discovered many barriers to implementation that must be resolved before payers invest in full-scale adoption [5]. This paper systematically reviews the research literature to identify observed barriers to remote health implementation, adoption, and retention for adult patients with type 2 diabetes in the United States. This review was restricted to a single country because of the substantial differences at the financial, strategic, operational, and tactical levels of health care systems and settings in different countries. Payment structures and technology access are 2 such conditions that frequently vary between countries. The barriers identified in this systematic review will inform the design and implementation of future remote health interventions for diabetes self-management in the United States. Researchers in other countries will also be more informed about barriers to remote health for diabetes from the results of this review, although the prevalence of each barrier category will undoubtedly vary from one country to another.

Literature Review

This review identified 24 other relevant reviews. Although these reviews provided no systematic analysis of barriers, it is important to compare and contrast their contributions with ours.

Reviews focused on diabetes outcomes for the general population reported improved health for patients using telemedicine as compared with those with regular care [6-10], whereas those focused on outcomes for specific populations saw mixed results [11-13]. Baig et al [11] addressed African American and Hispanic diabetic patients and concluded that rigorous evaluation of existing and new interventions was badly needed. Van den Berg et al [12] found that most elderly patients living at home were able to use telemedicine devices to good effects, whereas Sutcliffe et al [13] found no conclusive evidence that communication technologies improved health outcomes for younger patients.

Other reviews considered specific types of technology such as mobile phones and apps [14-16]. Baron et al [14] found methodological weaknesses and inconsistent evidence, whereas Holtz and Lauckner [15] saw positive trends resulting from mobile phone use. Chomutare et al [16] identified the most prevalent features of mobile phone diabetes apps and found personal education to be the most underrepresented. Arnhold et al [17] evaluated mobile phone app usability and found apps with analysis functions to have the least favorable usability scores. Three reviews focused on telehealth for the broader area of chronic disease management. Dennis et al [18] found that telephone coaching improved health behaviors, self-efficacy, and health status, especially for vulnerable populations with limited access to health services. Hamine et al [19] found mixed evidence and called for more research on overcoming barriers, and van den Berg et al [12] found insufficient customization in telemedicine apps for older adult populations, a barrier identified in our study as well. Two reviews investigated how information technology affects diabetes self-management. El-Gayar et al [20] reported the need for more comprehensive, user-centered interventions, whereas Cotter et al [21] found Internet interventions to be viable options for diabetes self-management, especially those with personalized feedback, tracking, and peer support.

Two reviews analyzed behavioral telehealth interventions for glycemic control. Cassimatis and Kavanagh [4] found that behavioral telehealth interventions show promise in improving diabetes self-care and glycemic control, especially those emphasizing physical activity and dietary adherence. Behavioral change techniques such as feedback on performance, education on consequences, and self-monitoring were linked to positive changes in health behaviors, psychological well-being, and clinical parameters [3]. Two other reviews explored the cost of telemedicine for diabetes self-management and found little evidence to support claims of cost-effectiveness [1,22]. Greenwood et al [23] concluded that telehealth interventions rarely include all the elements of the care protocols recommended by the International Diabetes Federation. Lepard et al [24] compared telehealth with face-to-face interventions for rural adults with type 2 diabetes and found collaborative goal-setting to be effective for both.

Wilson et al [25] examined the barriers and facilitators of access to self-monitoring for minority populations. Cultural awareness, social expectations, and language were identified as barriers. Our study did not reveal these same barriers to be as prevalent, presumably due to reporting bias from the provider's perspective as opposed to the patient's perspective. In addition, each study design only included patients who could fluently speak the language used by the remote health intervention. Finally, a review by Radhakrishnan et al [26] identified barriers and facilitators for sustainability of telehomecare programs for chronic disease management, including barriers regarding health literacy of the patients and cost-effectiveness of remote health.

Some of these same barriers were identified in our search as well.

In this research, we develop an inventory of common barriers to the implementation of remote health interventions over the last 5 years. In addition, we analyze the frequency of each barrier type, outcome measures, terminology, and technology used. We also discuss dropout rates, payments to patients, reported costs of the remote health interventions, and suggestions to overcome barriers. Finally, we draw connections between our barrier inventory and a newly proposed conceptual model of necessary conditions for successful remote health.

Objective

The objective of this systematic review was to identify and classify barriers to remote health interventions for adult patients with type 2 diabetes in the United States. We define a *barrier* as any cause of reduced technology effectiveness.

Methods

This section describes retrieval procedures, inclusion and exclusion criteria, and the data collection process of this systematic review.

Retrieval Procedures

We searched the following 6 electronic databases: MEDLINE (Ovid), Embase (Ovid), CINAHL, Cochrane Central, Northern Light Life Sciences Conference Abstracts, and Scopus (Elsevier). The focus was on 2 primary concepts that were searched using keywords and appropriate index terms for each database: type 2 diabetes (diabetes mellitus, type 2, type II) and remote health (telemonitor, telemedicine, telehealth, mHealth, eHealth). Several variations of the term “noninsulin dependent” were included because this is another way to describe type 2 diabetes. Multimedia Appendix 1 provides specific queries for Medline, Embase, and CINAHL databases. Results were limited to papers published in English from 2010 to 2015. Articles were exported into Refworks (ProQuest) and duplicates removed. We also used Scopus (Elsevier) to search references and citations of included studies.

Inclusion and Exclusion Criteria

Each article was screened by 2 independent authors in Refworks, first by the title and abstract, then by the full text. To be included, a study had to (1) involve patients with type 2 diabetes; (2) show results of an implemented study or trial; (3) use remote health (or the terms mHealth, eHealth, telemonitor, telehealth, or telemedicine); (4) occur in the United States; and (5) use adult populations. We excluded all duplicate texts and articles without available full texts. In addition, articles that were editorials, systematic reviews, article reviews, personal opinion articles, case studies, and all other summary-type or synthesis-type articles were excluded [27]. Any discrepancies between the 2 reviewers were settled by a third reviewer or group consensus.

Data Collection Process and Data

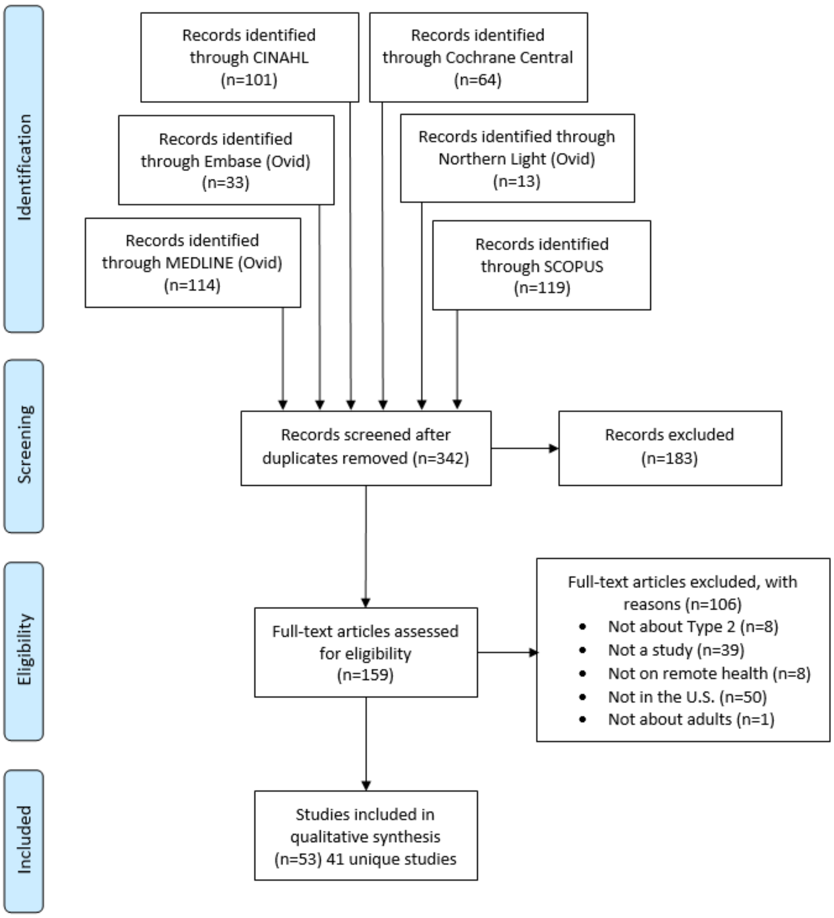
Reviewers worked in teams of 2 to code data from 53 papers using Google Forms. The form was created by 3 authors and then tested and revised by all authors prior to use. After the finalization of the Google Form, the 2 reviewers worked together to code each paper, discussing and agreeing on all the information before recording. Information was collected on study characteristics (time, length, location, type, facility, gender, participant age, socioeconomic status, eligibility, number of participants, costs, languages used), outcome measures, dropout rates, key results, remote health intervention name, technology used, barriers to implementation, and suggestions to reduce or eliminate the barriers.

Results

Results of the Study Selection

Figure 1 shows the study selection process. The initial search resulted in 342 citations from 6 databases. The original searches yielded 114 articles from Medline, 33 from Embase, 101 from CINAHL, 64 from Cochrane Central, 13 from Northern Light, and 119 from Scopus. There were 159 articles that passed title and abstract screening. From these, 106 articles were eliminated during full-text screening. The reasons for exclusion of the articles were as follows: 8 were not about type 2 diabetes, 39 were not studies, 8 were not about remote health, 50 were not in the United States, and 1 was not about adults. There were 53 articles that implemented remote health interventions for self-management of adult type 2 diabetes in the United States published from 2010 to 2015. Some of the publications were on the same project (IDEATel, TExT-MED, and Beacon Communities), so information for those articles was combined for a total of 41 unique studies that met search criteria. See Multimedia Appendix 2 for a summary of these articles and reports on findings for a limited number of categories (eg, outcome measures, technology used, barriers).

Figure 1. Selection process.



Study Characteristics

This section describes the characteristics of studies included in our systematic review. The following subsections analyze the types of studies, length of studies, study dates, participant characteristics (age, gender), study locations, languages, comorbidities, types of technology, terminology, outcome measures, and costs.

Types of Studies

Of the included studies, 51% (21/41) were randomized controlled trials (RCT), 24% (10/41) were pilot studies of a newly developed remote health intervention, and 10% (4/41) were quasi-experimental studies (Multimedia Appendix 3). One study [28] identified itself as a quasi-experimental pilot study, and thus was classified as both in our analysis. The “Other” category had the following entries: longitudinal study, nonrandomized parallel control group study, observational study, cluster-randomized clinical trial, prospective randomized trial, nonrandomized prospective observational preintervention or postintervention studies, and prospective longitudinal randomized trial.

Length of Study

Study length ranged from 2 weeks [29] to 5 years [30]. As shown in the Multimedia Appendix 3, % (21/41) of the papers were shorter than 6 months and 34% (14/41) of the studies lasted between 7 and 12 months. Only 12% (5/41) studies lasted more than a 1-year period. One study did not report its length [31,32]. The longest studies were 3.5 years [33] and 5 years [30].

Study Dates

All articles were published from 2010 to 2015, as required by the search criteria. There were 83% (34/41) studies that provided specific dates in which the study occurred. Of the studies that reported this information, all occurred from 2000 to 2013, with the earliest occurring during 2000 to 2005 [30].

Participant Characteristics

Number of Participants

The number of participants in the study varied from 11 [34] to 1838 [35,36] (Multimedia Appendix 3). The majority of studies (71%, 29/41) had less than 200 participants, and 10% (4/41) had more than 400 participants [30,35-38]. One study did not report the number of participants [39].

Participant Age

All studies were conducted on patients above 18 years of age (as specified in the inclusion criteria). One study reported the mean participant age in the 30s [40], 12% (5/41) studies in the 40s, 46% (19/41) in the 50s, and 20% (8/41) in the 60s (Multimedia Appendix 3). A single study had participants with mean age in the 70s [30]. There were 17% (7/41) studies [33,38,39,41-44] that did not report the mean participant age. Three studies [33,41,42] only provided participant age range.

Gender

Most studies (76%, 32/41) had a female to male ratio ranging from 38:62 [45] to 81:19 [33]. Stone et al [44] had only male participants, whereas 2 other studies [34,40] were mostly female. There were 12% (5/41) Veterans Affairs (VA) studies that had less than 7% female participants. There were 7% (3/41) of the selected studies [39,42,43] that did not report gender.

Location

The selected studies were analyzed by the geographic location in the United States using regions defined by the US Census Bureau [46]. After analyzing the selected studies by geographic regions, more studies (37%, 15/41) were conducted in the Midwest than any other region.

Multimedia Appendix 3 indicates that 22% (11/41) studies were conducted in the west, 32% (13/41) studies were conducted in the south, and 22% (9/41) studies were conducted in the northeast region. Note that the numbers in Multimedia Appendix 3 sum to 48 because some studies were in multiple states, for example, study [39] had locations in 3 states, viz, UT, LA, and MI.

Language

Most of the communications in the studies were exclusively English-based (71%, 29/41). A single study had a primary language other than English and was conducted in Spanish [33]. However, a considerable portion of the studies (22%, 9/41) were implemented with both English and Spanish options available for patients. Two studies [47,48] were available in English, Spanish, and Cantonese. No other languages were incorporated as part of the studies observed.

Comorbidity

The majority of papers (74%, 31/41) focused solely on type 2 diabetes. Others studied diabetes patients with comorbidities, including 12% (5/41) with hypertension and 7% (3/41) with cardiovascular conditions. Patients in the study by Henderson et al [42] exhibited comorbidities in hypertension, hyperlipidemia, and cardiovascular disease; patients in Dang et al [28] had comorbid hypertension and osteoarthritis; and patients in Abebe et al [30] had comorbid cognitive decline. One study included both type 1 and type 2 diabetes [49]. Multimedia Appendix 3 summarizes these findings.

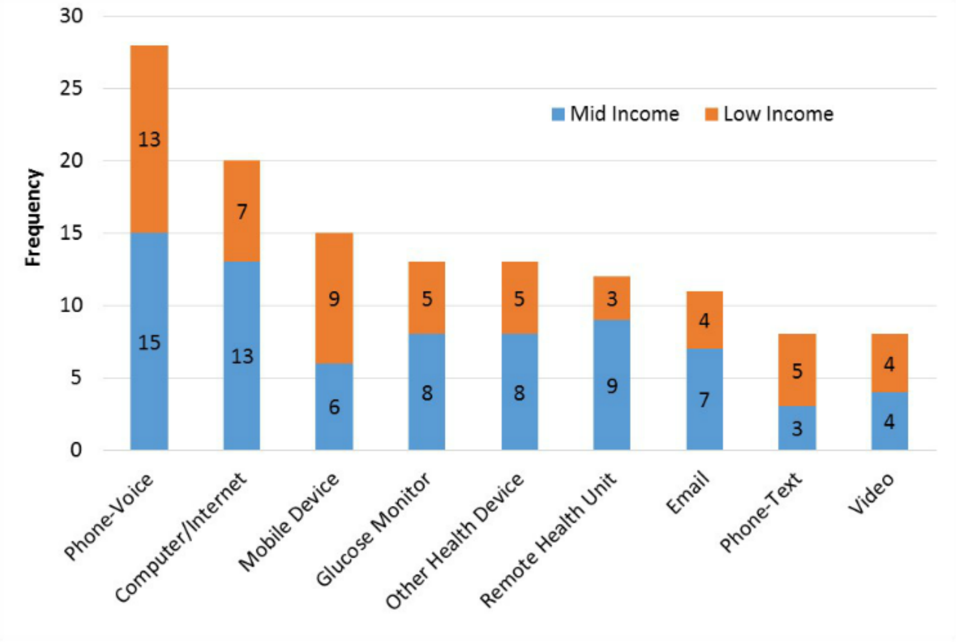
Technology

Studies used between 1-8 different types of technology (median=3) for remote health interventions. The following 9 technology categories were identified: Phone-Voice, Phone-Text, Mobile Device-Internet or Apps, Video, Email (computer or mobile phone), Remote Health Unit, Computer or Internet, Glucose Monitor, and “Other Health Device.” Phone-Voice indicates that phone calls were made via landlines or mobile phones. Phone-Text refers to the use of text messages as a communication method between the researchers and the participants. Mobile Device-Internet or Apps indicates the study allowed the participant to access health data through Internet or an app via mobile phones, iPads, or tablets.

Video involved the use of videoconferencing, but could be done through a mobile phone or tablet app or a laptop’s integrated camera, a desktop computer with a camera, or a remote health intervention with videoconferencing capabilities. Email was created as a category because many studies mentioned the use of email as a communication form but did not specify how it was to be accessed (eg, mobile device, laptop, desktop). Remote health units were single devices, often developed commercially, that combined different types of technology for both monitoring and communication. The Computer or Internet category indicates data were sent over the Internet using a computer, laptop, or other device. “Other Health Device” indicates the use of a health monitoring device, other than a blood glucose monitor, such as a blood pressure monitor or telemetry device. Glucose monitors were counted as a unique category as they are highly associated with diabetes self-management.

Figure 2 shows the utilization of different types of technology used in the low- and mid-income studies. Phone-Voice was used by 68% (28/41) of the studies and Computer or Internet was used by 49% (20/41) of the studies. Also, 37% (15/41) of the studies mentioned using a mobile device for Internet or app access. There were 29% (12/41) studies that used a remote health unit like the Authentidate Electronic House Call [50]. Only 5 low-income studies used a blood glucose monitor compared with 8 regular income studies. Finally, 20% (8/41) studies used video conferencing in their program and 27% (11/41) used email as a source of communication.

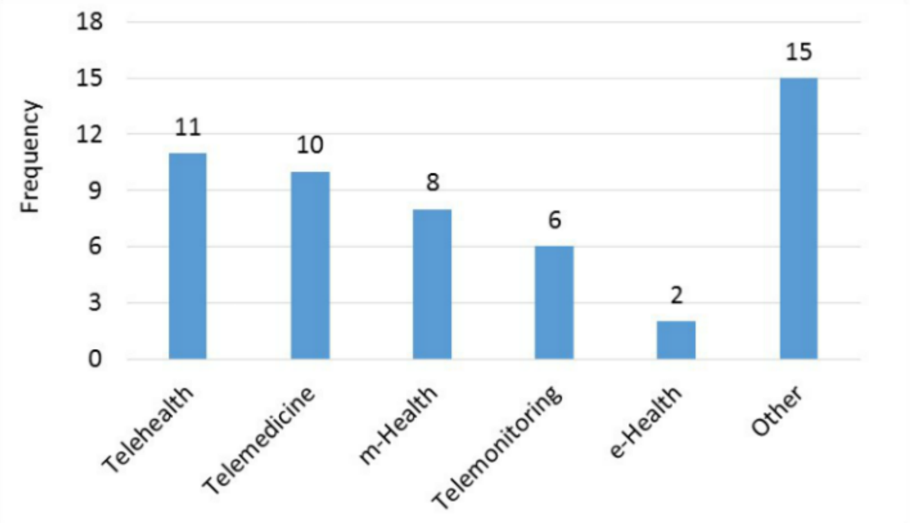
Figure 2. Technology used by popularity.



Terminology

There was a variety of terminology used to describe the remote health intervention used in each study (Figure 3). There were a total of 20 unique terms. The most popular term found in our articles was “telehealth” which was used in 27% (11/41) of the studies. This was followed by “telemedicine” which was mentioned in 24% (10/41) of the articles and “m-Health or mHealth” was the primary term for 20% (8/41) of the studies. However, there were 15 terms in the “other” category because each was only found in a single paper. These “other” terms were: telecommunications, televisits, text-messaging intervention, voice intervention, automated telephone, Web-based, medical assisted coaching (MAC), automated telephone self-management support program, home health monitoring, text-message based program, mobile phone personalized behavioral intervention, teleconsultation, diabetes care telemonitoring device trial, remote monitoring, and telephone care management.

Figure 3. Terminology popularity index.



Outcome Measures

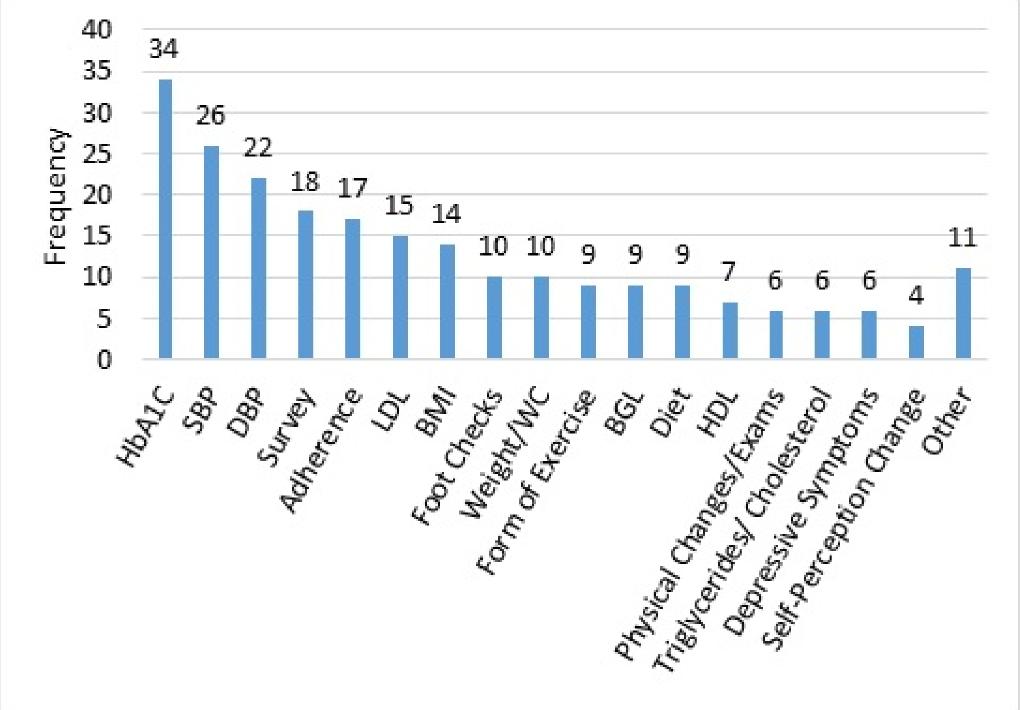
Each study used between 1 and 16 outcomes measures (median=5) to evaluate the effectiveness of the remote health intervention. The outcome measures varied across studies. The commonly used outcome measures are tallied in Figure 4. HbA1C was the most prevalent outcome measure and 83% (34/41) of the papers utilized this measure. Davis et al [51] referred to glycated hemoglobin as GHb and used it to evaluate the effectiveness of telehealth diabetes self-management, but we considered this as the same term as HbA1C. Blood glucose level was an outcome measure for 22% (9/41) of the papers. All the papers that measured the sugar level also included HbA1C counts except Aikens et al [52].

There were 63% (26/41) of the studies that measured blood pressure. The 11 of 12 low-income studies measured systolic blood pressure (SBP) and diastolic blood pressure (DBP). But of the 14 mid-income papers that checked SBP, only 11 checked DBP [34,53,54]. High-density lipoprotein (HDL) was not as frequently seen as an outcome measure as low-density lipoprotein (LDL). HDL is popularly known as the "good cholesterol" and physicians are more concerned with regulating the high-level of LDL or "bad cholesterol" in patients' blood. For example, of the 37% (15/41) papers that recorded LDL, only 7 also measured HDL outcomes. Body mass index (BMI) was reported as an outcome measure in 34% (14/41) of the papers. Figure 4 combines waist circumference and weight into a single category which was used in 24% (10/41) of the studies.

Self-efficacy or adherence either in diet, exercise, medication, or diabetes management was an outcome measure in 41% (17/41) of the studies. This was labeled in Figure 4 as "Adherence." For instance, there was self-efficacy in Diabetes Management Practices Scale [55]; adherence in immunizations [40]; the Medication Adherence Self-Efficacy Scale [29,56]; and adherence to the use of home telemedicine unit, home monitoring, and diabetes health maintenance [30]. Physical Changes or Exams (labeled "Physical Exams" in Figure 4) were performed in 15% (6/41) of the papers. The type of physical exam varied greatly from the Framingham Risk Score [28], Albumin-to-creatinine ratio [51], glycemic control [57], and Charlson Comorbidity Index [30]. None of the mid-income studies had Physical Changes or Exams as outcome measures as opposed to 6 low-income studies. Furthermore, 24% (10/41) of the studies included foot checks, and 22% (9/41) checked for diet and exercise levels.

There were 43% (18/41) of the studies that used surveys, interviews, and mixed methods to gather information on participants. Some studies [29,39,57] implemented surveys to understand the patient's satisfaction rate with the programs. Others sought to gage the participant's knowledge level on diabetes medication [29]. Ruggiero et al [58] implemented a survey to understand the environmental barriers participants experienced that prevented patients from taking their medication. Finally, it should be noted that there were less frequently used "other" outcome measures noted in 27% (11/41) of the studies that gathered information on weekly activities, Self-monitoring of blood glucose (SMBG) measurement data, fat intake, fructosamine blood level, dilated eye exam, diabetes distress, microalbumin, health literacy, trigger safety concerns, alcohol intake, and smoking status, such as Palmas et al [30] and Pressman et al [45].

Figure 4. Outcome measures.



Cost

In the review, 88% (36/41) of the studies did not mention cost estimates of the remote health intervention. Davis et al [51] and Henderson et al [42] described a method to compute costs but did not report results. Only Palmas et al [30] documented costs, estimating a total of US $622 per person/month. This cost was US $358/month from technology vendors, US $115/month from the bioinformatics team, and US $149/month from the clinical teams. However, this estimate did not account for participant's transportation savings due to the remote health intervention. Fischer et al [38] reported an overall direct programmatic cost of US $134,750 for the 20-month intervention. Greenwood et al [23] reported US $34.20 per participant for phone and messaging over 9 months, and Katalenich et al [59] varied depending on the intervention; however, these ranged from US $12.52 to US $15.50 per person.

Main Results

In this section, we define 5 categories of barriers, catalogue and inventory of each barrier, discuss barrier impacts, and present a conceptual model of conditions under which remote health can be successful and sustainable.

Barrier Classification

After identifying barriers to successful remote health interventions among the included studies, the authors then organized the barriers into 5 categories based on their relation to a resource (*patient, provider, health system, digital infrastructure, or intervention design*). *Patient (P)* barriers are the causes of reduced effectiveness originating with the patient, such as health illiteracy. *Technology Access (T)* barriers arise from restricted availability of supporting digital infrastructure, such as poor wireless coverage. *Design (D)* barriers are shortcomings in the specified technical configuration and intended mode of use, such as inadequate support for provider feedback. *Provider (Pv)* barriers are those originating with the provider's care team and environment, such as staff training, and *System (S)* barriers include health care organization factors external to the care environment that affect technology effectiveness, such as limited institutional support. All identified barriers fell within these categories. Furthermore, barrier prevalence was related to income level and thus this information was recorded. Table 1 provides an inventory of barriers identified, organized by these types, along with income level.

Table 1. Barrier inventory.

Barrier type		Income level		
		Low	Mid	Total
Patient barriers				33
P1	Low formal education	4	0	4
P2	Technology illiteracy (uncomfortable with technology)	7	3	10
P3	Medication nonadherence	3	2	5
P4	Patients desire in-person contact with provider (perceived lack of confidence and comfort)	3	0	3
P5	Low perceived value or effectiveness	2	2	4
P6	Health illiteracy	4	1	5
P7	Other	1	1	2
Technology access barriers				21
T1	Patient does not have required technology	5	3	8
T2	Technology is cost prohibitive to the patient (not affordable)	4	1	5
T3	Limited internet access in the area	3	0	3
T4	Other	3	2	5
Design barriers				60
D1	Lack of customization to patient preferences and needs	5	3	8
D2	Lack of accuracy or reliability (patient or provider)	7	6	13
D3	Content not engaging or relevant	3	6	9
D4	Timing of patient-provider interactions	2	1	3
D5	Decisions of content and frequency of interventions	3	3	6
D6	Patients not incorporated into the design needs	3	0	3
D7	No analysis on impact with comorbidities	2	1	3
D8	Labor- and time-intensive for providers	4	2	6
D9	Other	4	5	9
Provider barriers				14
Pv1	Data accessibility to patient logs (access to patient logs)	2	1	3
Pv2	Low integration into provider work flow	3	1	4
Pv3	Other	3	4	7
System barriers				20
S1	Limitations on scalability	1	9	10
S2	Lack of program reimbursement by insurance	1	2	3
S3	High cost of intervention	1	2	3
S4	Other	3	1	4

Barrier Inventory

Each of the 41 included studies reported or showed evidence of 0-12 barriers (median=3, mean=3.6). Three studies [33,45,60] discussed limitations of the study designs, but did not report barriers associated with design and implementation of the remote health intervention.

Patient Barriers

Patient Barriers (P1-P7) were identified in 33 studies. Technology illiteracy (uncomfortable with technology) was reported in 24% (10/41) of the studies, whereas low health literacy and low formal education were reported in 12% (5/41) and 10% (4/41) of the studies, respectively. Several authors noted that VA populations tended to be more technology savvy than general populations [52,54,61-63]. In many instances, patients could use computers and the Internet, but felt uncomfortable or lacked confidence in Internet-based communication and preferred using phone calls.

Technology Access Barriers

Technology Access Barriers (T1-T4) were identified 21 times. The main barrier (20%, 8/41) was that the patient did not own or have access to the required technology. For instance, Arora et al [49] eliminated 51 of 74 potential candidates because they did not have a text-capable mobile phone. In some cases, the patient could not afford to purchase the technology (12%, 5/41). Consequently, studies either limited participation to patients able to afford the technology or it was provided. It was noted that some patients started with the technology but eventually dropped out due to costs. In other instances, the patient had the technology available but had limited access to the Internet (7%, 3/41).

Design Barriers

Design Barriers (D1-D9) were noted 61 times. The most common *Design Barrier* was inaccurate or unreliable data (32%, 13/41), as most studies required patients to manually enter their own data and were therefore prone to recall bias or human error. A total of 22% (9/41) studies reported issues with the content of the intervention not being engaging or relevant, and 20% (8/41) reported a lack of customization of the intervention. This was particularly true for studies that utilized manual input or did not provide language alternatives for nonnative English speakers.

Provider Barriers

Provider Barriers (Pv1-Pv3) were referenced in 14 studies, the most common being poor integration of remote health technology and provider work flow (10%, 4/41). There were physician complaints of the system being intrusive [53,64] and a single study described the work flow challenges arising from the call center not being located near the clinical areas [41]. Personnel shortage and insufficient training were also mentioned, as was limited transparency of patient health data [53,64]. The latter most commonly occurred in systems relying on manual data collection and uploads where the patient failed to provide the data at the time of communication with the providers.

System Barriers

System Barriers (S1-S4) were reported 21 times with the most common being limitations on scalability in 24% (10/41) studies. Scalability issues were most commonly cited because the study involved only a few providers or had specialized populations (eg VA patients), and the authors were unsure of how the results would apply to larger, more general populations. There were 6 S *ystem Barriers* classified as “Other” because they were less common and included high cost of interventions [30], undiversified population of VA patients [28], unreported cost-effectiveness [57], and an uncontrolled study design [49].

Most barriers occurred in both low- and mid-income populations, but some were more prevalent in one population than the other. For example, *Patient Barriers* were more prevalent in low-income studies which had 24 instances compared with the mid-income studies that had only 9 instances, with low-income patients having more difficulty using the technology. *System Barriers* were most prominent in mid-income studies (14 times) as compared with low-income (6 times) with limitations on scalability to larger population sizes being the most prevalent barrier.

Barrier Impacts

Dropout Rates

The patient dropout rate was a commonly reported problem ranging from 5% to 57% and averaging 22%. Buis et al [35,36] experienced the largest at 57%. There were 27% (11/41) studies that did not report a dropout rate [23,28,34,38,39,42,44,59,62,65,66]. Multimedia Appendix 3 provides a histogram of dropout rates. There were 14% (6/41) studies with dropout rates 10% or lower, and 3 (7%, 3/41) with dropout rates greater than 50% [30,35,36,67]. A regression analysis revealed that there is no strong correlation (R^2=.0639) between the study period and dropout rates. The next section compares dropout rates with payments.

Payments and Dropout Rates

There were 26% (11/41) studies that paid individuals for participation and reported dropout rates (Figure 5). Eight of these were low-income populations and 4 exceeded the average dropout rate of 19%. Interestingly, the 3 highest paying studies [40,48,68] were all low-income and were above the 19% average dropout rate. Anderson et al [41] (low-income) compensated participants with US $25 (17% dropout) and Wakefield et al [53] gave a US $20 gift card for their cooperation at the end of the study (23% dropout). Katz et al [40] (low-income) relied heavily on texting glucose readings to physicians and gave active participants US $20 monthly cellphone waiver for 1 year if they entered glucose readings that month (50% dropout). Arora et al [68] (low-income) paid participants US $175 for participating 6 continuous months (28% dropout) and assisted 3 people by awarding US $20 monthly stipends to alleviate costs of upgrading mobile phone service to unlimited texting capability.

Arora et al [49] (low-income) paid US $50 after successful completion of a 3-week study (13% dropout). Heisler et al [56] (low-income) gave participants US $20 stipend per assessment, which consisted of taking HbA1C readings (6% dropout). The study by Grilo et al [69], which lasted a period of 6 months, paid US $10 for completing 3 months in the program (18% dropout). Davis et al [51] (low-income) gave a gift card for each completed clinical visit, but did not specify the gift card's monetary amount (18% dropout). Dick et al [70] (mid-income) gave their participants US $25 to cover the expenses of unlimited text messaging plan and US $30 for their participation (5% dropout). Aikens et al [61] (low-income) gave US $20 for patients at baseline then another US $20 at study completion of the 3-6 month study (13% dropout). Finally, Ratanawongsa et al [48] (low-income) awarded individuals a US $25 gift card per participation and US $50 for each interview spaced 6 months apart over a 1-year period (23% dropout) up to US $150.

Payments amounts do not appear to have a direct correlation with dropout rates, but this observation may be confounded by the fact that payment amounts are intended for different purposes in each study (eg travel, technology) and are a reflection of the effort required. For example, the highest paying study required a 2-year commitment for regularly entering glucometer readings.

There were 46% (19/41) studies that did not pay for participation that also reported the dropout rate (Figure 6). The average dropout rate for nonpaid participation studies was 23% compared with 19% for paid participation studies.

Figure 5. Dropout rates for paid participation studies.

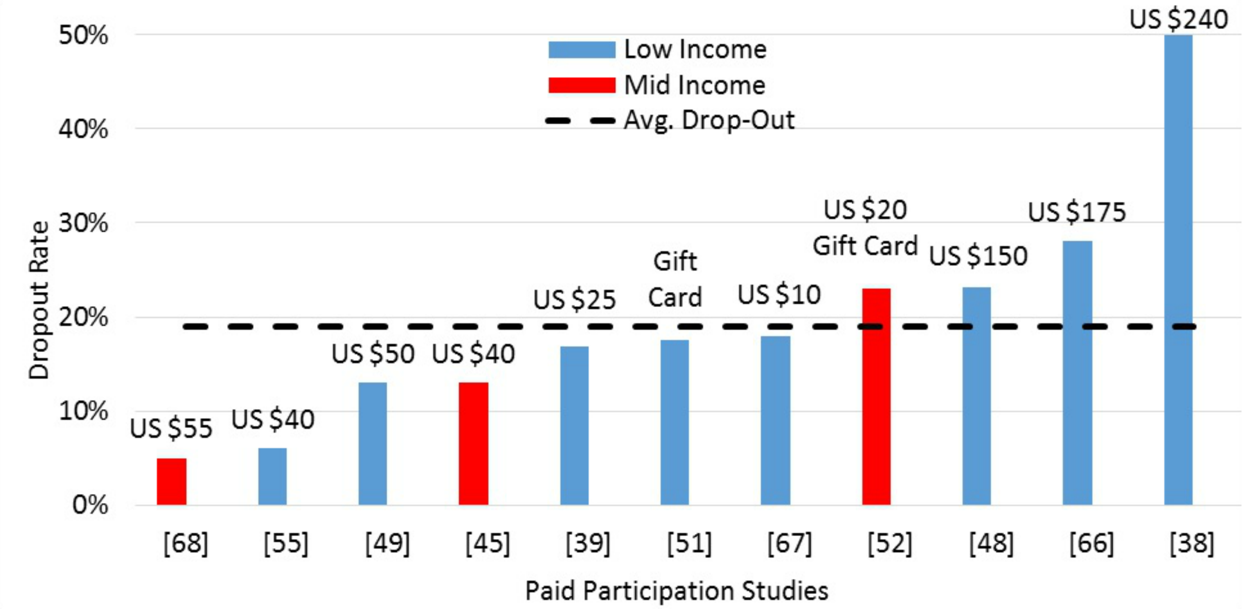
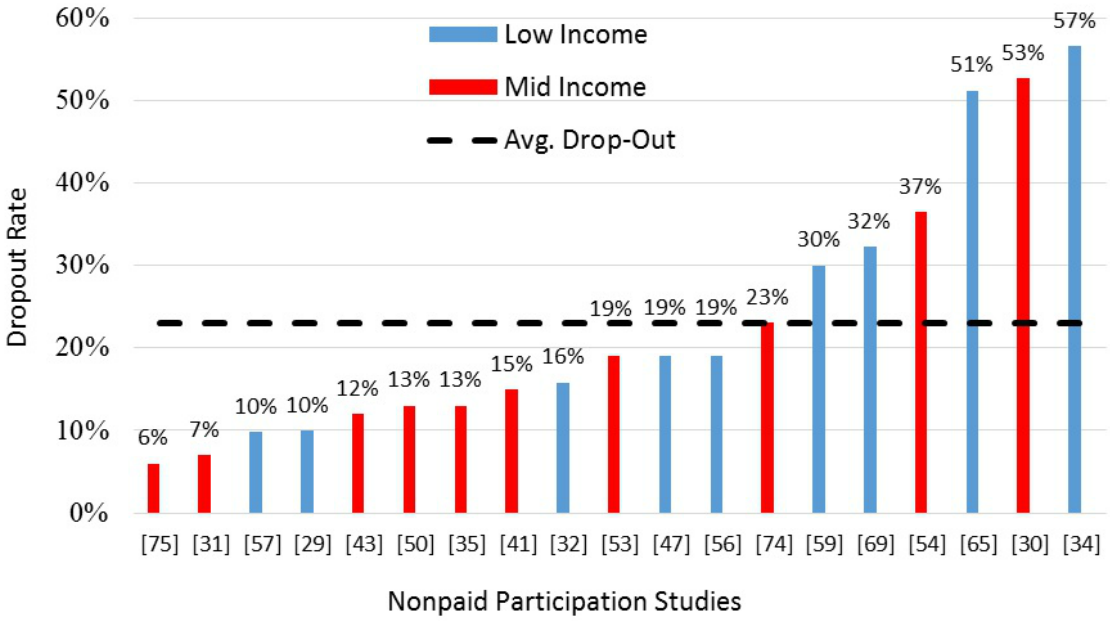


Figure 6. Dropout rates for nonpaid participation studies.



Barriers and Dropout Rates

There was no strong relationship between high dropout rates and the number of barriers identified in each study. We suspect this lack of correlation is due to the journal articles not explicitly reporting on barriers and the low sample size of papers. However, we did note that studies with higher than average dropout rates had an average of 4.6 barriers, compared with those studies with lower than average dropout rates having an average of only 3.5 barriers.

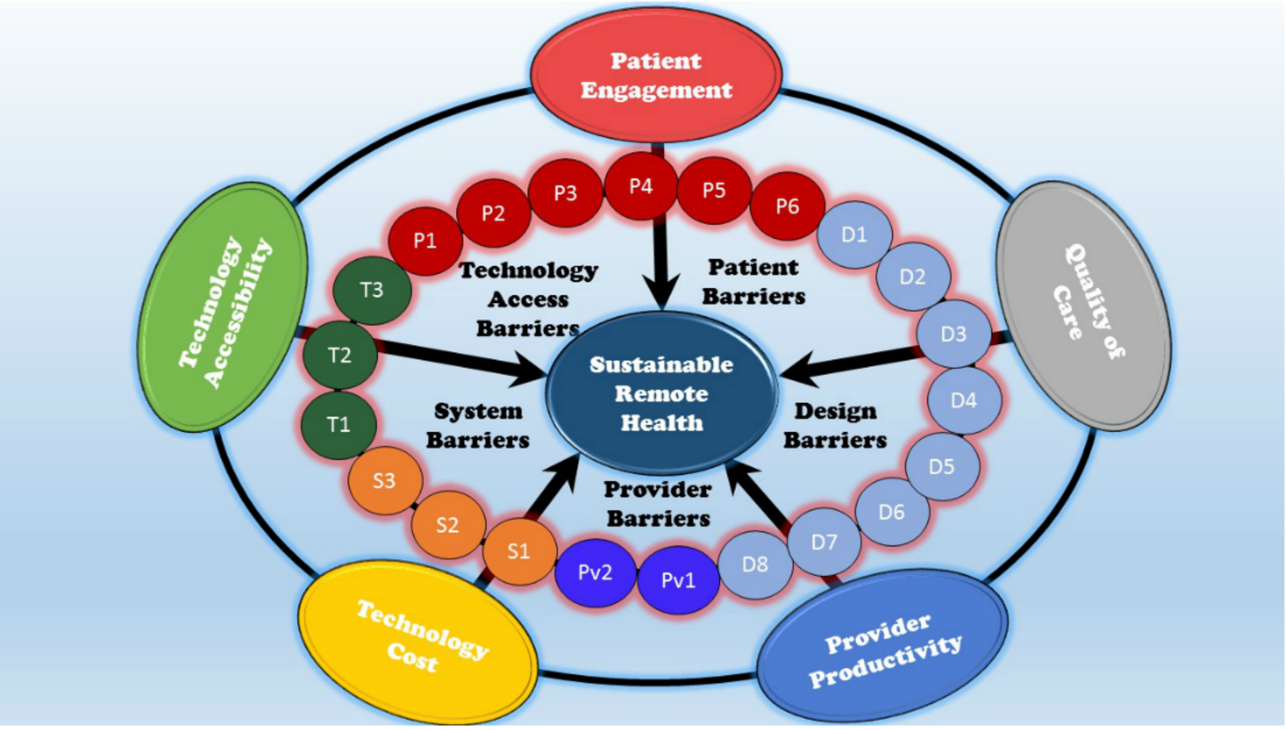
Suggestions to Reduce or Eliminate Barriers

Declining engagement or participation (a symptom of barriers) was the most commonly cited problem. Some studies reported that the intervention was time-intensive [40] and others simply attributed it to poor follow-up [33]. However, no data were recorded on why patients did not continue. Tang et al [37] and Heisler et al [56] proposed that personalized strategies to achieve health goals (eg, taking medication) might keep patients engaged. Studies with minority populations [71] suggested providing culturally tailored educational activities into the patient's daily life, so that interventions occur at multiple levels and are conveniently delivered. A few studies suggested that declining participation could be attributed to the cost of maintaining the technology (eg, cell phone bills). One study involving mobile phone messaging [68] helped participants pay for unlimited text messaging, thereby eliminating the patient's participation cost. But the dropout rate remained high at 28%. Carter et al [55] equipped patients with laptops containing wireless broadband cards for this same reason, but again dropout rates were high at 37%. To improve adherence to the new technology, more training and involvement is needed while communicating with patients [29,40].

Conceptual Model

The triple aim in health care is often referred to as *access, quality, and cost* [72]. There is increasing recognition that *patient-engagement* is critical to achieving high-quality, affordable care [73,74] due to the prevalence of chronic diseases that are behavior- and lifestyle-driven. Moreover, systems theory has shown that system *productivity* is a requirement to obtain quality control at an affordable cost [75]. *Provider productivity* is especially important in an era of aging population, growing chronic disease, and decreasing primary care resources. We used these 5 elements to develop a conceptual model (Figure 7) for effective remote health, adapting it around technology and organizing it around the 2 main players in the system, patients, and provider. We included insurance companies in the provider side for parsimony in the model. We then mapped the barriers into these elements. The mapping is not one-to-one, as some barriers affect multiple elements. We believe these elements are 5 necessary conditions for successful and sustainable remote health. Although we do not argue that the list is complete or that meeting these conditions is sufficient for sustainable success, we do believe that many of the challenges of remote health would begin to dissolve if these conditions were achieved. We now discuss how the barriers identified in this review affect these 5 elements.

Figure 7. Conceptual model of barriers to successful, sustainable remote health.



Patient Engagement

The remote health intervention of the future must ensure regular *patient engagement* to promote positive behavior, efficacy, and retention. Designing a remote health intervention with strong patient engagement will favorably affect patient dropout rates. Unfortunately, this category had the largest number of barrier instances with 60 identified (Barriers T3, P1-P6, D1, D3, D5, and D6) in the systematic review. Barriers to patient engagement include low formal education, technology illiteracy, and health illiteracy on the part of the patient, as well as low perceived value of the system, a preference for personal contact with their provider, and medication nonadherence. Also, lack of Internet accessibility in the patient's area, lack of customization of the system design, poor content choices, inconvenient frequencies of interventions, and lack of inclusion in the system design are also barriers to high levels of patient engagement. Clearly, if patients cannot or will not engage in remote care, remote health systems will fail.

Quality of Care

It is important not to compromise on quality of care delivered by the remote health intervention. The system must provide high *quality of care* that leads to improved health outcomes, improved perceptions of care, and higher efficacy. There were 28 instances (Barriers D2, D4, D5, D7, and Pv1) of barriers that threaten the ability to provide patients with quality care, which include lack of accuracy of data (eg, manual input), failure to understand the impact of remote health interventions on comorbid conditions, timing, and frequency of intervention and patient-provider–interactions, and accessibility of patient data records to the provider. Again, if quality of care cannot be assured, remote health will not succeed.

Provider Productivity

The remote health intervention must improve *provider productivity* so that a diminishing number of providers can care for an increasing number of patients. There were 23 instances (Barriers D4, D8, Pv2, and S1) of provider productivity barriers identified in the systematic review. Remote health interventions that are labor-intensive, require a high number of patient-provider–interactions, have poor integration into the provider work flow, and are limited in scalability contain barriers to increased provider productivity and cannot be sustained.

Technology Accessibility

Remote health interventions must be *accessible* for the patient. Accessibility implies technology that is affordable to the patient and ubiquitous at the point of use to enhance adoption and retention. The systematic review identified 16 instances (Barriers T1-T3) in which technology accessibility to the patient was a barrier because the patient did not already have the technology, could not afford the technology, or had limited Internet access. Technology accessibility must be achieved for remote health to have an impact.

Technology Cost

The remote health intervention will need to be *cost-effective* for long-term sustainability. This implies an overall reduction in health care spending, due to reduced emergency department and hospitalization, which exceeds the capital costs providers must pay to implement and maintain the remote system. The systematic review identified 16 instances (Barriers S1-S3) of barriers to cost-effective technology, which include limitations on scalability, lack of program reimbursement by insurance, and high-cost of interventions. Remote health systems that do not pay for themselves will not be sustainable.

Risk of Bias Within and Across Studies

All the articles included in the systematic review had their own internal biases. Mainly, remote health approaches rely heavily on the data submitted by the patients [67,70,76]. These data can be willingly distorted or abstained by patients. Other patients may have the best self-reporting intentions, but could have been confused on the proper way to use the remote health intervention [34,59]. In addition, the outcomes were not always reliable. For example, Bell et al [77] counted the number of diabetes videos watched by patients each month. However, there is a big difference between "watching" a video and actually "paying attention" and "understanding" a video. Other studies had selection bias, especially all the VA studies because the conclusions for this population group are not necessarily generalizable to the rest of the US population. Many were written with an optimistic perspective on remote health and did not take a critical viewpoint in addressing the challenges encountered during implementation. Even if the barriers were noted during a study, the authors of the study may not have reported them in the journal article as they were not the focus of the publication or due to space limitations of the journal. Furthermore, the barriers identified in certain populations, such as patient barriers in mid-income populations, could be attributed to reporting bias by the researchers who assume patients are having difficulty adopting to technology instead of examining possible system or provider issues. Finally, many of the selected articles were pilot studies with 47% (20/41) having population sizes less than 100. In addition, 54% (22/41) of the studies were shorter than 6 months. Thus, the barriers identified are not those that would necessarily arise during large-scale, long-term implementation of a remote health intervention for type 2 diabetes self-management.

Discussion

Limitations

The authors recognize several limitations in this systematic review. Although several databases were searched, including databases with conference proceedings, as well as references of included studies, it is possible some articles fitting the inclusion criteria were missed. There are several descriptive terms for remote health and the authors do not claim that the search criteria included a comprehensive list of these terms. Identifying barriers was especially challenging as they were not always explicitly labeled and the authors had to use their best judgment to identify, interpret, and classify the barriers in each study. When coding the articles, the authors may have overlooked some barriers because of the manner in which they were described in the article. In some instances, the authors had a difficult time agreeing whether some issues were in fact barriers, or a symptom of a barrier (eg, patient engagement); thus, the identification and determination of barriers were subjective and could have been overlooked by a particular reviewer. The final list of barriers collected is not necessarily comprehensive but was created based on findings from the included studies. Finally, for study characteristic results, some studies only had limited information available such as not specifying the type of blood pressure (systolic or diastolic) used or not specifying the device used for communication via email and Internet.

Conclusions

This systematic review analyzed papers with remote health interventions for type 2 diabetes patients. Technology in remote health reduces the burden of diabetes by providing patients with medical resources and education without the need to leave their homes. Six databases were searched for peer-reviewed journal articles published between 2010 and 2015 that implemented a remote health intervention for type 2 diabetes care. A total of 53 papers were selected on 41 different studies. The principal findings of this systematic review included analysis on barrier classification and inventory. Lack of data accuracy was the most common barrier as it was identified in 32% (13/41) of all studies and was equally found in both low- and mid-income populations. The lack of data accuracy was often a result of manual reporting or input of monitoring data. Concerns over scalability were cited in 24% (10/41) of studies (mostly mid-income) and technology illiteracy was observed in 24% (10/24) of studies (mostly low-income). Declining patient engagement was observed in almost every study as a result of these barriers.

Few studies addressed mechanisms for reducing barriers. For those that did, suggestions were made to customize the strategies or provide culturally tailored solutions to increase patient engagement. Some studies thought that assistance with technology education or cost would reduce dropout rates. Around 29% (12/41) of the studies paid patients for participation. Some reviews maintain that there is insufficient evidence to conclude that remote health interventions significantly improved type 2 diabetes outcomes [13]. The total cost was only reported in a single study, thus leaving open questions on cost-effectiveness of remote health. A review by Radhakrishnan et al [26] was most similar to ours because it identified barriers and facilitators for sustainability of telehomecare programs for chronic disease management. Although some barriers were common to both (including health literacy and cost-effectiveness), our review identifies many additional barriers, including scalability, provider training, and system design.

We also proposed a conceptual model for successful implementation of remote health interventions. The model explains that technology accessibility, increased patient engagement, technology cost, increased provider productivity, and increased quality of care are 5 necessary conditions for remote health. Focusing on the barriers that impede these necessary conditions (eg, technology illiteracy and data accuracy) will better connect the patients to the clinics and providers for successful implementation of a remote health intervention for diabetes self-management in the United States. The results of this systematic review will facilitate other research in the design of remote health technology interventions as we identify common impediments in the design, implementation, adoption, and communication of remote health for diabetes patients. Specifically, technology must advance to improve reporting accuracy and reliability of the data communicated from the patient to the provider. Identifying ways to address the scalability of remote health interventions should also be a priority, as well as innovative designs that allow customized interventions and increase patient engagement. The barrier inventory provides visibility and evidence that these are the most dominant, pressing challenges facing the advancement of remote health today.

Acknowledgments

This research has been partially supported by the Center for Remote Health Technologies and Systems at Texas A&M University.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms used for Medline and CINAHL.

[PDF File (Adobe PDF File), 176KB - jmir_v19i2e28_app1.pdf]

Multimedia Appendix 2

Summary of 41 studies in systematic review showing the author names, income levels of patient participants, terminology used, study type, technology used, outcome measures, and barriers.

[PDF File (Adobe PDF File), 67KB - jmir_v19i2e28_app2.pdf]

Multimedia Appendix 3

Statistics for study characteristics.

[PDF File (Adobe PDF File), 303KB - jmir_v19i2e28_app3.pdf]

References

1. Zhai Y, Zhu W, Cai Y, Sun D, Zhao J. Clinical- and cost-effectiveness of telemedicine in type 2 diabetes mellitus: a systematic review and meta-analysis. Medicine (Baltimore) 2014 Dec;93(28):e312 [FREE Full text] [doi: 10.1097/MD.0000000000000312] [Medline: 25526482]
2. Siriwardena LS, Wickramasinghe WA, Perera KL, Marasinghe RB, Katulanda P, Hewapathirana R. A review of telemedicine interventions in diabetes care. J Telemed Telecare 2012 Apr;18(3):164-168. [doi: 10.1258/jtt.2012.SFT110] [Medline: 22362832]
3. van Vugt M, de Wit M, Cleijne WH, Snoek FJ. Use of behavioral change techniques in web-based self-management programs for type 2 diabetes patients: systematic review. J Med Internet Res 2013 Dec 13;15(12):e279 [FREE Full text] [doi: 10.2196/jmir.2800] [Medline: 24334230]
4. Cassimatis M, Kavanagh DJ. Effects of type 2 diabetes behavioural telehealth interventions on glycaemic control and adherence: a systematic review. J Telemed Telecare 2012 Dec;18(8):447-450. [doi: 10.1258/jtt.2012.GTH105] [Medline: 23209266]
5. Klonoff DC. The current status of mHealth for diabetes: will it be the next big thing? J Diabetes Sci Technol 2013 May 01;7(3):749-758 [FREE Full text] [Medline: 23759409]
6. Verhoeven F, van Gemert-Pijnen L, Dijkstra K, Nijland N, Seydel E, Steehouder M. The contribution of teleconsultation and videoconferencing to diabetes care: a systematic literature review. J Med Internet Res 2007 Dec 14;9(5):e37 [FREE Full text] [doi: 10.2196/jmir.9.5.e37] [Medline: 18093904]
7. Verhoeven F, Tanja-Dijkstra K, Nijland N, Eysenbach G, van Gemert-Pijnen L. Asynchronous and synchronous teleconsultation for diabetes care: a systematic literature review. J Diabetes Sci Technol 2010 May 01;4(3):666-684 [FREE Full text] [Medline: 20513335]
8. Mignerat M, Lapointe L, Vedel I. Using telecare for diabetic patients: a mixed systematic review. Health Policy Technol 2014 Jun;3(2):90-112. [doi: 10.1016/j.hlpt.2014.01.004]
9. Marcolino MS, Maia JX, Alkmim MB, Boersma E, Ribeiro AL. Telemedicine application in the care of diabetes patients: systematic review and meta-analysis. PLoS One 2013;8(11):e79246 [FREE Full text] [doi: 10.1371/journal.pone.0079246] [Medline: 24250826]
10. Huang Z, Tao H, Meng Q, Jing L. Management of endocrine disease. Effects of telecare intervention on glycemic control in type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. Eur J Endocrinol 2015 Mar;172(3):R93-101 [FREE Full text] [doi: 10.1530/EJE-14-0441] [Medline: 25227131]
11. Baig AA, Wilkes AE, Davis AM, Peek ME, Huang ES, Bell DS, et al. The use of quality improvement and health information technology approaches to improve diabetes outcomes in African American and Hispanic patients. Med Care Res Rev 2010 Oct;67(5 Suppl):163S-197S [FREE Full text] [doi: 10.1177/1077558710374621] [Medline: 20675350]
12. van den Berg N, Schumann M, Kraft K, Hoffmann W. Telemedicine and telecare for older patients--a systematic review. Maturitas 2012 Oct;73(2):94-114. [doi: 10.1016/j.maturitas.2012.06.010] [Medline: 22809497]
13. Sutcliffe P, Martin S, Sturt J, Powell J, Griffiths F, Adams A, et al. Systematic review of communication technologies to promote access and engagement of young people with diabetes into healthcare. BMC Endocr Disord 2011 Jan 06;11:1 [FREE Full text] [doi: 10.1186/1472-6823-11-1] [Medline: 21210964]
14. Baron J, McBain H, Newman S. The impact of mobile monitoring technologies on glycosylated hemoglobin in diabetes: a systematic review. J Diabetes Sci Technol 2012 Sep;6(5):1185-1196 [FREE Full text] [Medline: 23063046]
15. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
16. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011 Sep 22;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
17. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
18. Dennis SM, Harris M, Lloyd J, Powell DG, Faruqi N, Zwar N. Do people with existing chronic conditions benefit from telephone coaching? A rapid review. Aust Health Rev 2013 Jun;37(3):381-388. [doi: 10.1071/AH13005] [Medline: 23701944]
19. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
20. El-Gayar O, Timsina P, Nawar N, Eid W. A systematic review of IT for diabetes self-management: are we there yet? Int J Med Inform 2013 Aug;82(8):637-652. [doi: 10.1016/j.ijmedinf.2013.05.006] [Medline: 23792137]
21. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: a systematic review of the evidence. J Diabetes Complications 2014;28(2):243-251 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.07.003] [Medline: 24332469]
22. Fitzner KK, Heckinger E, Tulas KM, Specker J, McKoy J. Telehealth technologies: changing the way we deliver efficacious and cost-effective diabetes self-management education. J Health Care Poor Underserved 2014 Nov;25(4):1853-1897. [doi: 10.1353/hpu.2014.0157] [Medline: 25418248]

23. Greenwood DA, Hankins AI, Parise CA, Spier V, Olveda J, Buss KA. A comparison of in-person, telephone, and secure messaging for type 2 diabetes self-management support. Diabetes Educ 2014 Apr 17;40(4):516-525. [doi: 10.1177/0145721714531337] [Medline: 24742540]
24. Lepard MG, Joseph AL, Agne AA, Cherrington AL. Diabetes self-management interventions for adults with type 2 diabetes living in rural areas: a systematic literature review. Curr Diab Rep 2015 Jun;15(6):608. [doi: 10.1007/s11892-015-0608-3] [Medline: 25948497]
25. Wilson C, Alam R, Latif S, Knighting K, Williamson S, Beaver K. Patient access to healthcare services and optimisation of self-management for ethnic minority populations living with diabetes: a systematic review. Health Soc Care Community 2012 Jan;20(1):1-19. [doi: 10.1111/j.1365-2524.2011.01017.x] [Medline: 21749529]
26. Radhakrishnan K, Xie B, Berkley A, Kim M. Barriers and facilitators for sustainability of tele-homecare programs: a systematic review. Health Serv Res 2016 Feb;51(1):48-75. [doi: 10.1111/1475-6773.12327] [Medline: 26119048]
27. Greenhalgh T. Getting your bearings: What is this paper about? In: How to read a paper: the basics of evidence-based medicine. 5th ed. UK: John Wiley & Sons Inc; Mar 1, 2014:28-44.
28. Dang S, Sanchez A, Oropesa L, Roos BA, Florez H. Telehealth-assisted care coordination of older veterans with type 2 diabetes lowers coronary heart disease risk despite clinical inertia. Diabetes Technol Ther 2010 Dec;12(12):995-1001. [doi: 10.1089/dia.2010.0066] [Medline: 21128846]
29. Osborn CY, Mulvaney SA. Development and feasibility of a text messaging and interactive voice response intervention for low-income, diverse adults with type 2 diabetes mellitus. J Diabetes Sci Technol 2013 May;7(3):612-622 [FREE Full text] [Medline: 23759393]
30. Palmas W, Shea S, Starren J, Teresi JA, Ganz ML, Burton TM, IDEATel Consortium. Medicare payments, healthcare service use, and telemedicine implementation costs in a randomized trial comparing telemedicine case management with usual care in medically underserved participants with diabetes mellitus (IDEATel). J Am Med Inform Assoc 2010;17(2):196-202 [FREE Full text] [doi: 10.1136/jamia.2009.002592] [Medline: 20190064]
31. Toledo FG, Ruppert K, Huber KA, Siminerio LM. Efficacy of the telemedicine for reach, education, access, and treatment (TREAT) model for diabetes care. Diabetes Care 2014 Aug;37(8):e179-e180. [doi: 10.2337/dc13-1909] [Medline: 25061149]
32. Siminerio L, Ruppert K, Huber K, Toledo FG. Telemedicine for reach, education, access, and treatment (TREAT): Linking telemedicine with diabetes self-management education to improve care in rural communities. Diabetes Educ 2014;40(6):797-805. [doi: 10.1177/0145721714551993] [Medline: 25253624]
33. Mayes PA, Silvers A, Prendergast JJ. New direction for enhancing quality in diabetes care: utilizing telecommunications and paraprofessional outreach workers backed by an expert medical team. Telemed J E Health 2010 Apr;16(3):358-363 [FREE Full text] [doi: 10.1089/tmj.2009.0110] [Medline: 20406123]
34. Bartlett A, Bosken L, Clark S, Glaser L, Stankovski S. The effects of a combined aerobic and resistive exercise program with the use of telehealth monitoring in reducing the value of cardiovascular risk factors of individuals vulnerable to developing type 2 diabetes. Cardiopulm Phys Ther J 2011;22(4):39-40.
35. Buis LR, Hirzel L, Turske SA, Des Jardins TR, Yarandi H, Bondurant P. Use of a text message program to raise type 2 diabetes risk awareness and promote health behavior change (part I): assessment of participant reach and adoption. J Med Internet Res 2013;15(12):e281 [FREE Full text] [doi: 10.2196/jmir.2928] [Medline: 24356329]
36. Buis LR, Hirzel L, Turske S, Des Jardins T, Yarandi H, Bondurant P. Use of a text message program to raise type 2 diabetes risk awareness and promote health behavior change (part II): Assessment of participants' perceptions on efficacy. J Med Internet Res 2013;15:e282. [doi: 10.2196/jmir.2929] [Medline: 24356359]
37. Tang PC, Overhage JM, Chan AS, Brown NL, Aghighi B, Entwistle MP, et al. Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. J Am Med Inform Assoc 2013 May 1;20(3):526-534 [FREE Full text] [doi: 10.1136/amiajnl-2012-001263] [Medline: 23171659]
38. Fischer HH, Eisert SL, Everhart RM, Durfee MJ, Moore SL, Soria S, et al. Nurse-run, telephone-based outreach to improve lipids in people with diabetes. Am J Manag Care 2012 Feb;18(2):77-84 [FREE Full text] [Medline: 22435835]
39. Abebe NA, Capozza KL, Des Jardins TR, Kulick DA, Rein AL, Schachter AA, et al. Considerations for community-based mHealth initiatives: insights from three Beacon Communities. J Med Internet Res 2013;15(10):e221 [FREE Full text] [doi: 10.2196/jmir.2803] [Medline: 24128406]
40. Katz R, Mesfin T, Barr K. Lessons from a community-based mHealth diabetes self-management program: “it's not just about the cell phone”. J Health Commun 2012;17 Suppl 1:67-72. [doi: 10.1080/10810730.2012.650613] [Medline: 22548601]
41. Anderson DR, Christison-Lagay J, Villagra V, Liu H, Dziura J. Managing the space between visits: a randomized trial of disease management for diabetes in a community health center. J Gen Intern Med 2010 Oct;25(10):1116-1122 [FREE Full text] [doi: 10.1007/s11606-010-1419-5] [Medline: 20556536]
42. Henderson JA, Chubak J, O'Connell J, Ramos MC, Jensen J, Jobe JB. Design of a randomized controlled trial of a web-based intervention to reduce cardiovascular disease risk factors among remote reservation-dwelling American Indian adults with type 2 diabetes. J Prim Prev 2012 Aug;33(4):209-222 [FREE Full text] [doi: 10.1007/s10935-012-0276-x] [Medline: 23001642]

43. Stone RA, Rao RH, Sevick MA, Cheng C, Hough LJ, Macpherson DS, et al. Active care management supported by home telemonitoring in veterans with type 2 diabetes: the DiaTel randomized controlled trial. Diabetes Care 2010 Mar;33(3):478-484 [FREE Full text] [doi: 10.2337/dc09-1012] [Medline: 20009091]
44. Stone RA, Sevick MA, Rao RH, Macpherson DS, Cheng C, Kim S, et al. The diabetes telemonitoring study extension: an exploratory randomized comparison of alternative interventions to maintain glycemic control after withdrawal of diabetes home telemonitoring. J Am Med Inform Assoc 2012;19(6):973-979 [FREE Full text] [doi: 10.1136/amiajnl-2012-000815] [Medline: 22610495]
45. Pressman AR, Kinoshita L, Kirk S, Barbosa GM, Chou C, Minkoff J. A novel telemonitoring device for improving diabetes control: protocol and results from a randomized clinical trial. Telemed J E Health 2014 Feb;20(2):109-114. [doi: 10.1089/tmj.2013.0157] [Medline: 24404816]
46. Census. 2010. Census regions and divisions of the United States URL: http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf [accessed 2017-01-05] [WebCite Cache ID 6nHTrFxI6]
47. Lyles CR, Schillinger D, Lopez A, Handley M, Ratanawongsa N, Sarkar U. Safety events during an automated telephone self-management support intervention. J Diabetes Sci Technol 2013 May;7(3):596-601 [FREE Full text] [Medline: 23759391]
48. Ratanawongsa N, Handley MA, Quan J, Sarkar U, Pfeifer K, Soria C, et al. Quasi-experimental trial of diabetes self-management automated and real-time telephonic support (SMARTSteps) in a medicaid managed care plan: study protocol. BMC Health Serv Res 2012;12:22 [FREE Full text] [doi: 10.1186/1472-6963-12-22] [Medline: 22280514]
49. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. Diabetes Technol Ther 2012 Jun;14(6):492-496. [doi: 10.1089/dia.2011.0252] [Medline: 22524591]
50. Shane-McWhorter L, Lenert L, Petersen M, Woolsey S, McAdam-Marx C, Coursey JM, et al. The utah remote monitoring project: improving health care one patient at a time. Diabetes Technol Ther 2014 Oct;16(10):653-660 [FREE Full text] [doi: 10.1089/dia.2014.0045] [Medline: 24991923]
51. Davis RM, Hitch AD, Salaam MM, Herman WH, Zimmer-Galler IE, Mayer-Davis EJ. TeleHealth improves diabetes self-management in an underserved community: diabetes TeleCare. Diabetes Care 2010 Aug;33(8):1712-1717 [FREE Full text] [doi: 10.2337/dc09-1919] [Medline: 20484125]
52. Aikens JE, Zivin K, Trivedi R, Piette JD. Diabetes self-management support using mHealth and enhanced informal caregiving. J Diabetes Complications 2014;28(2):171-176 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.11.008] [Medline: 24374137]
53. Wakefield BJ, Koopman RJ, Keplinger LE, Bomar M, Bernt B, Johanning JL, et al. Effect of home telemonitoring on glycemic and blood pressure control in primary care clinic patients with diabetes. Telemed J E Health 2014 Mar;20(3):199-205 [FREE Full text] [doi: 10.1089/tmj.2013.0151] [Medline: 24404819]
54. Wakefield BJ, Holman JE, Ray A, Scherubel M, Adams MR, Hillis SL, et al. Effectiveness of home telehealth in comorbid diabetes and hypertension: a randomized, controlled trial. Telemed J E Health 2011 May;17(4):254-261. [doi: 10.1089/tmj.2010.0176] [Medline: 21476945]
55. Carter EL, Nunlee-Bland G, Callender C. A patient-centric, provider-assisted diabetes telehealth self-management intervention for urban minorities. Perspect Health Inf Manag 2011;8:1b [FREE Full text] [Medline: 21307985]
56. Heisler M, Choi H, Palmisano G, Mase R, Richardson C, Fagerlin A, et al. Comparison of community health worker-led diabetes medication decision-making support for low-income Latino and African American adults with diabetes using e-health tools versus print materials: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S13-S22 [FREE Full text] [doi: 10.7326/M13-3012] [Medline: 25402398]
57. Iannitto JM, Dickman K, Lakhani RH, So MJC. Telehealth insulin program: managing insulin in primary care. J Nurse Pract 2014 Sep;10(8):567-574. [doi: 10.1016/j.nurpra.2014.07.027]
58. Ruggiero L, Moadsiri A, Quinn LT, Riley BB, Danielson KK, Monahan C, et al. Diabetes island: preliminary impact of a virtual world self-care educational intervention for african americans with type 2 diabetes. JMIR Serious Games 2014;2(2):e10 [FREE Full text] [doi: 10.2196/games.3260] [Medline: 25584346]
59. Katalenich B, Shi L, Liu S, Shao H, McDuffie R, Carpio G, et al. Evaluation of a remote monitoring system for diabetes control. Clin Ther 2015 Jun 1;37(6):1216-1225 [FREE Full text] [doi: 10.1016/j.clinthera.2015.03.022] [Medline: 25869625]
60. Leichter SB, Bowman K, Adkins RA, Jelsovsky Z. Impact of remote management of diabetes via computer: the 360 study--a proof-of-concept randomized trial. Diabetes Technol Ther 2013 May;15(5):434-438. [doi: 10.1089/dia.2012.0323] [Medline: 23537419]
61. Aikens JE, Rosland A, Piette JD. Improvements in illness self-management and psychological distress associated with telemonitoring support for adults with diabetes. Prim Care Diabetes 2015 Apr;9(2):127-134 [FREE Full text] [doi: 10.1016/j.pcd.2014.06.003] [Medline: 25065270]
62. McFarland M, Davis K, Wallace J, Wan J, Cassidy R, Morgan T, et al. Use of home telehealth monitoring with active medication therapy management by clinical pharmacists in veterans with poorly controlled type 2 diabetes mellitus. Pharmacotherapy 2012 May;32(5):420-426. [doi: 10.1002/j.1875-9114.2011.01038.x] [Medline: 22488512]
63. Wakefield BJ, Holman JE, Ray A, Scherubel M, Adams MR, Hills SL, et al. Outcomes of a home telehealth intervention for patients with diabetes and hypertension. Telemed J E Health 2012 Oct;18(8):575-579. [doi: 10.1089/tmj.2011.0237] [Medline: 22873700]

64. Koopman RJ, Wakefield BJ, Johanning JL, Keplinger LE, Kruse RL, Bomar M, et al. Implementing home blood glucose and blood pressure telemonitoring in primary care practices for patients with diabetes: lessons learned. Telemed J E Health 2014 Mar;20(3):253-260 [FREE Full text] [doi: 10.1089/tmj.2013.0188] [Medline: 24350806]
65. Toledo FG, Triola A, Ruppert K, Siminerio LM. Telemedicine consultations: an alternative model to increase access to diabetes specialist care in underserved rural communities. JMIR Res Protoc 2012 Nov;1(2):e14 [FREE Full text] [doi: 10.2196/resprot.2235] [Medline: 23612044]
66. McMahon GT, Fonda SJ, Gomes HE, Alexis G, Conlin PR. A randomized comparison of online- and telephone-based care management with internet training alone in adult patients with poorly controlled type 2 diabetes. Diabetes Technol Ther 2012 Nov;14(11):1060-1067 [FREE Full text] [doi: 10.1089/dia.2012.0137] [Medline: 22953754]
67. Stamp KD, Allen NA, Lehrer S, Zagarins SE, Welch G. Telehealth program for medicaid patients with type 2 diabetes lowers hemoglobin A1c. J Manag Care Med 2012;15:39-46.
68. Arora S, Peters AL, Burner E, Lam CN, Menchine M. Trial to examine text message-based mHealth in emergency department patients with diabetes (TExT-MED): a randomized controlled trial. Ann Emerg Med 2014 Jun;63(6):745-54.e6. [doi: 10.1016/j.annemergmed.2013.10.012] [Medline: 24225332]
69. Grilo SA, Shallcross AJ, Ogedegbe G, Odedosu T, Levy N, Lehrer S, et al. Food insecurity and effectiveness of behavioral interventions to reduce blood pressure, New York City, 2012-2013. Prev Chronic Dis 2015;12(E16):1-7 [FREE Full text] [doi: 10.5888/pcd12.140368] [Medline: 25674675]
70. Dick JJ, Nundy S, Solomon MC, Bishop KN, Chin MH, Peek ME. Feasibility and usability of a text message-based program for diabetes self-management in an urban African-American population. J Diabetes Sci Technol 2011 Sep;5(5):1246-1254 [FREE Full text] [Medline: 22027326]
71. Ruggiero L, Riley BB, Hernandez R, Quinn LT, Gerber BS, Castillo A, et al. Medical assistant coaching to support diabetes self-care among low-income racial/ethnic minority populations: randomized controlled trial. West J Nurs Res 2014 Oct;36(9):1052-1073 [FREE Full text] [doi: 10.1177/0193945914522862] [Medline: 24569698]
72. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769 [FREE Full text] [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
73. Clancy CM. Patient engagement in health care. Health Serv Res 2011 Apr;46(2):389-393 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01254.x] [Medline: 21371026]
74. Barello S, Graffigna G, Vegni E. Patient engagement as an emerging challenge for healthcare services: mapping the literature. Nurs Res Pract 2012;2012:1-7 [FREE Full text] [doi: 10.1155/2012/905934] [Medline: 23213497]
75. Deming W. Out of the Crisis. Cambridge, MA: MIT Press; 1986.
76. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
77. Bell AM, Fonda SJ, Walker MS, Schmidt V, Vigersky RA. Mobile phone-based video messages for diabetes self-care support. J Diabetes Sci Technol 2012;6(2):310-319 [FREE Full text] [Medline: 22538140]

Abbreviations

BMI: body mass index
CINAHL: Current Index to Nursing and Allied Health Literature
DBP: diastolic blood pressure
eHealth: electronic health
EHR: electronic health record
GHb: glycated hemoglobin
HbA1C: glycated hemoglobin
HDL: high-density lipoprotein
IDEATel: Informatics for Diabetes and Education Telemedicine
LDL: low-density lipoprotein
MAC: Medical Assistant Coaching
mHealth: mobile health
PHQ-9: Patient Health Questionnaire Depressive Screen
RCT: Randomized Control Trial
SBP: systolic Blood Pressure
SMS: short message service
TExT-MED: text message–based mobile health intervention

Edited by CL Parra-Calderón; submitted 22.07.16; peer-reviewed by M Harris, B Holtz; comments to author 20.09.16; revised version received 05.11.16; accepted 23.11.16; published 13.02.17.

Please cite as:
Alvarado MM, Kum HC, Gonzalez Coronado K, Foster MJ, Ortega P, Lawley MA
Barriers to Remote Health Interventions for Type 2 Diabetes: A Systematic Review and Proposed Classification Scheme
J Med Internet Res 2017;19(2):e28
URL: http://www.jmir.org/2017/2/e28/
doi:10.2196/jmir.6382
PMID:28193598

Original Paper

From the Experience of Interactivity and Entertainment to Lower Intention to Smoke: A Randomized Controlled Trial and Path Analysis of a Web-Based Smoking Prevention Program for Adolescents

Georges Elias Khalil[1], MPH, PhD; Hua Wang[2], PhD; Karen Sue Calabro[1], DrPH; Natasha Mitra[3], BA; Ross Shegog[4], PhD; Alexander V Prokhorov[1], MD, PhD

[1]MD Anderson Cancer Center, Department of Behavioral Sciences, University of Texas, Houston, TX, United States
[2]University at Buffalo, Department of Communication, Buffalo, NY, United States
[3]Health Science Center at San Antonio, School of Public Health, University of Texas, San Antonio, TX, United States
[4]Health Science Center at Houston, Center for Health Promotion and Prevention Research, University of Texas, Houston, TX, United States

Corresponding Author:
Alexander V Prokhorov, MD, PhD
MD Anderson Cancer Center
Department of Behavioral Sciences
University of Texas
1155 Pressler St. Unit 1330
Houston, TX, 77030
United States
Phone: 1 7137452382
Fax: 1 7137454286
Email: aprokhor@mdanderson.org

Abstract

Background: Web-based programs for smoking prevention are being increasingly used with some success among adolescents. However, little is known about the mechanisms that link the experience of such programs to intended nicotine or tobacco control outcomes.

Objective: Based on the experiential learning theory and extended elaboration likelihood model, this study aimed to evaluate the impact of a Web-based intervention, A Smoking Prevention Interactive Experience (ASPIRE), on adolescents' intention to smoke, while considering the experience of interactivity and entertainment as predictors of reduced intention to smoke, under a transitional user experience model.

Methods: A total of 101 adolescents were recruited from after-school programs, provided consent, screened, and randomized in a single-blinded format to 1 of 2 conditions: the full ASPIRE program as the experimental condition (n=50) or an online , text-based version of ASPIRE as the control condition (n=51). Data were collected at baseline and immediate follow-up. Repeated-measures mixed-effects models and path analyses were conducted.

Results: A total of 82 participants completed the study and were included in the analysis. Participants in the experimental condition were more likely to show a decrease in their intention to smoke than those in the control condition (beta=−0.18, *P*=.008). Perceived interactivity (beta=−0.27, *P*=.004) and entertainment (beta=−0.20, *P*=.04) were each associated with a decrease in intention to smoke independently. Results of path analyses indicated that perceived interactivity and perceived entertainment mediated the relationship between ASPIRE use and emotional involvement. Furthermore, perceived presence mediated the relationship between perceived interactivity and emotional involvement. There was a direct relationship between perceived entertainment and emotional involvement. Emotional involvement predicted a decrease in intention to smoke (beta=−0.16, *P*=.04).

Conclusions: Adolescents' experience of interactivity and entertainment contributed to the expected outcome of lower intention to smoke. Also, emphasis needs to be placed on the emotional experience during Web-based interventions in order to maximize reductions in smoking intentions. Going beyond mere evaluation of the effectiveness of a Web-based smoking prevention program, this study contributes to the understanding of adolescents' psychological experience and its effect on their intention to smoke.

With the results of this study, researchers can work to (1) enhance the experience of interactivity and entertainment and (2) amplify concepts of media effects (eg, presence and emotional involvement) in order to better reach health behavior outcomes.

Trial Registration: Clinicaltrials.gov NCT02469779; https://clinicaltrials.gov/ct2/show/NCT02469779 (Archived by WebCite at http://www.webcitation.org/6nxyZVOf0)

(J Med Internet Res 2017;19(2):e44) doi:10.2196/jmir.7174

KEYWORDS

smoking prevention; intention to smoke; adolescent; Web-based intervention; interactivity; entertainment; emotions; presence; experience

Introduction

Tobacco smoking remains the most preventable cause of death in the world [1]. However, more than 20% of adolescents in the United States have smoked at least one cigarette [2], and their use of new and emerging nicotine and tobacco products has increased [2-5]. Approximately 80% of adult daily tobacco smokers begin using before the age of 18 years and become addicted during adolescence [1,6]. This suggests the need for continuous efforts to prevent smoking among adolescents. In particular, Web-based programs are increasingly used to improve tobacco risk communication and prevent smoking [7-9].

While several Web-based programs have shown success in delaying smoking initiation among adolescents [10-14], little is known about adolescents' experience with such programs. This is evident from meta-analyses and reviews on the evaluation of smoking prevention programs [10-12,15]. First, evaluation studies tend to focus solely on *health outcomes* (eg, intention to smoke and smoking initiation behavior [10-12,15]) without examining psychological media effects that may explain change in outcomes. The investigation of media effects related to one's experience during program use may shed light on how a Web-based program achieves success. Second, cognitive theories dominate program evaluations [11,16], leaving aside the role of emotional involvement in driving successful outcomes. In particular, studying cognitive processes of behavior change alone can be limiting, especially with adolescents whose decision making is often dependent upon their emotions.

To fill in these gaps, it is important to better understand the processes by which Web-based interventions work to reach health outcomes. Going beyond a mere evaluation that explains "whether" a program works, it is crucial to investigate "how" a program works to be successful. This is done by studying the underlying mechanism that delineates one's experience of a Web-based intervention.

As a result, in this study, we aimed to test associations that link participation in a Web-based program to changes in a health outcome. We achieved this aim in the context of a Web-based intervention called *A Smoking Prevention Interactive Experience* (ASPIRE) [17-20]. Web-based interventions such as ASPIRE make use of two main design elements: interactivity and entertainment [20]. *Interactivity* refers to a technology design that allows for a two-way communication between adolescents and the program (eg, wide range of options, ability to explore virtual environments, and connect with other users) [21]. *Entertainment* refers to a design that is expected to drive enjoyment and gratification (eg, comedy, humor, fear appeal, and game play) [22-25]. In ASPIRE, features of interactivity and entertainment are applied through entertaining videos that intertwine with interactive and entertaining activities relevant to smoking-related issues. Several associations that link interactivity and entertainment to health outcomes deserve attention.

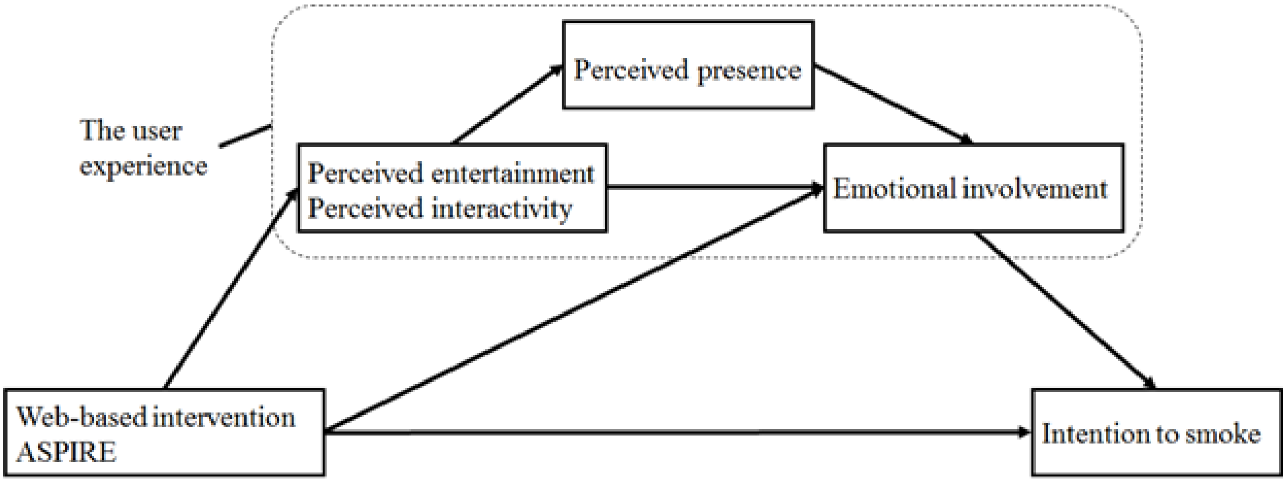
With a careful design of interactive and entertaining features, users' experience of interactivity and entertainment becomes crucial for the success of Web-based programs [26-30]. *Perceived interactivity* [31] refers to the user's perception of an active two-way communication with the program and a sense of control over online behavior [21,26-29,32]. *Perceived entertainment* [31,33] refers to the users' beliefs that their experience with the program is enjoyable and entertaining [30,34-36]. Both perceived interactivity and perceived entertainment are built on the idea that interactivity and entertainment are subjective experiences [26-29]. Research has supported a multidimensional perspective of entertainment that is based on the experience of psychological amusement, pleasant atmosphere, and joy [37]. Research has also supported that users must be psychologically involved and build a sense of influence in order to experience interactivity [21].

The effect of the experience of interactivity and entertainment on health outcomes is supported by the experiential learning theory (ELT) [38] and the extended elaboration likelihood model (E-ELM) [39]. According to the ELT, users of interactive applications are able to learn through the exploration of environments. In essence, first-hand exploration fosters curiosity and ultimately facilitates learning. According to the E-ELM, individuals exposed to entertainment programming begin to feel transported into the program environment by (1) becoming immersed in the world of the program and (2) becoming emotionally involved in the program. Ultimately, such transportation drives support for the healthy behavior.

Supportive of the ELT and E-ELM, empirical findings show that users' experience of interactivity and entertainment gradually drives them toward healthy outcomes. Perceived interactivity [40-44] and entertainment [22-25] have been found to drive attention to media platforms. This attention is expected to create *perceived presence* (ie, individuals' perception of being present inside the environment of the website [45]). Perceived presence is a key outcome in the context of Web-based interventions because interactivity and entertainment involve the experience of environments that demand attention and immersion. As a result of perceived presence, users of Web-based programs may experience *emotional involvement* (ie, the intensity of felt emotions as a result of the intervention [46-49]). Research has shown that perceived presence is greater in the context of emotional environments, and a greater emotional state is experienced as a result of a higher level of presence [50]. Ultimately, emotional involvement can have an impact on psychological antecedents of nicotine or tobacco use (eg, lowering intention to smoke) through affective persuasion [51]. Such line of research suggests a model by which adolescents move from intervention exposure to changes in tobacco control outcomes.

Although such associations have been examined each on its own, they remain fragmented. As a result, through a short-term randomized controlled trial with ASPIRE, this study developed a user experience model and statistically validated it using path analysis (Figure 1). We hypothesized the following: exposure to ASPIRE is positively related to perceived interactivity and perceived entertainment (hypothesis 1); adolescents who use ASPIRE are more likely to show a decrease in intention to smoke compared with adolescents in a control group (hypothesis 2); adolescents' perceived interactivity and perceived entertainment during ASPIRE are positively related to a decrease in intention to smoke (hypothesis 3); perceived interactivity and perceived entertainment mediate the relationship between ASPIRE use and emotional involvement (hypothesis 4); perceived presence mediates the relationship between perceived entertainment and emotional involvement (hypothesis 5); perceived presence in the ASPIRE environment mediates the relationship between perceived interactivity and emotional involvement in ASPIRE (hypothesis 6); and emotional involvement is related to a decrease in intention to smoke (hypothesis 7).

Figure 1. Conceptual model tested in this study. ASPIRE: A Smoking Prevention Interactive Experience.



Methods

Study Design

In order to examine the net benefit of interactivity and entertainment in ASPIRE, this study was conducted post hoc, using data from a 2-arm single-blinded randomized controlled trial with assessments at baseline and immediate follow-up (time × condition). This trial, called ASPIRE Reactions, was conducted in 2014. It is registered at the ClinicalTrials.gov registry (identifier: NCT02469779). Its components adhere to the CONSORT (Consolidated Standards of Reporting Trials) and CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines [52,53].

ASPIRE and ASPIRE-Control

The trial involved 2 conditions: (1) the ASPIRE program and (2) a control condition without interactivity or entertainment (ASPIRE-control). Study manipulations are presented in Table 1 (see Multimedia Appendix 1). ASPIRE was presented in 4 sessions, 40 minutes each, spread over a period of 1 month [54].

Table 1. Manipulations for study conditions.

Main factor and elements	ASPIRE[a]	ASPIRE-control
Interactivity		
Two-way communication	Yes	No
Control over platform	Yes	No
Clicking behavior	Yes	No
Virtual environments	Yes	No
Entertainment		
Narrative or storytelling	Yes	No
Music; sound effects	Yes	No
Virtual characters or avatars	Yes	No
Channel	Multimedia (video, audio, and text)	Text only
Content	Facts delivered in a personal context (age-tailored)	Facts only
Involvement	Emotional (peripheral processing of animation and narratives)	Cognitive (central processing of facts)

[a] ASPIRE: A Smoking Prevention Interactive Experience.

ASPIRE features interactivity and entertainment to engage adolescent users through text, animations, videos, and task-oriented activities (Multimedia Appendix 2). Text mainly involves educational facts about tobacco. As part of the videos, users are presented with cartoon animations and testimonies from high school students. ASPIRE's multimodality is also characterized by the presence of activities that allow users to explore health information in a two-dimensional environment. An activity in ASPIRE entails an environment with an expected task. By completing the task, the user gets to uncover health information. For instance, as part of an activity, users attempt to make a virtual character smoke in order to discover the poisons and chemicals that are left in the lungs (Multimedia Appendix 3). Also, ASPIRE allows users to enter a mediated environment in which they can search for clues, click on objects or characters, and listen to health messages [55]. In addition, ASPIRE includes videos that rely on humorous messages, dramatic events, and real-life stories about tobacco consequences. ASPIRE can be accessed online for free [54].

ASPIRE-control was designed to include the same health information presented in ASPIRE but without any features of interactivity or entertainment. In order to design an appropriate control condition, we conducted qualitative content analysis of ASPIRE’s content to identify factual information about tobacco or smoking that is depicted in ASPIRE. Factual information was compiled and summarized to form a text-based document. Following this procedure, the document was fragmented to form a series of paragraphs. The paragraphs presented information in the same order as presented in ASPIRE. In order to control for exposure to online health information, the text was introduced in a mock website that had the same background design as the actual ASPIRE intervention.

Sample Recruitment and Procedure

A total of 4 after-school programs located in Medicaid-eligible districts of Houston, Texas, were randomly selected for participant recruitment, including the Boys and Girls Clubs (2 sites), the Salvation Army Boys and Girls Clubs (1 site), and the Young Men's Christian Association (YMCA; 1 site). After approval from the program directors, a verbal announcement reached 509 adolescents. Interested adolescents completed child assent and parental permission. We assessed adolescent eligibility through a screening conducted before participation. Inclusionary criteria were as follows: being of ages 12 through 18 years, being a student in a middle school or high school, and being a nonsmoker (have not smoked in the past year, not even one cigarette, cigar, or hookah).

All participants in the final sample obtained parental consent. Recruitment and data collection took a period of 4 months. The institutional review boards for human subjects research at the University of Texas MD Anderson Cancer Center and the University at Buffalo, the State University of New York approved this study.

All participants started their experience with the intervention 3 days after they completed the baseline survey. The principal investigator generated the random allocation sequence. Research assistants assigned enrolled adolescents to groups. We used concealed envelopes to conduct randomization. Participants were not told which intervention was the intervention of interest. In ASPIRE and ASPIRE-control, participants used identical computers and had private space for individual viewing and headphones for noise reduction. A research assistant was available for technical assistance and supervision. At the end of the intervention, participants completed the follow-up survey. Then, 20 randomly selected participants from the ASPIRE group participated in exit interviews (data not included in this paper).

As an ethical consideration, after follow-up assessment, the ASPIRE-control group received information about the actual ASPIRE intervention and ways to access the website. Each participant was offered a US $15 gift card for participation in the study. Participants received giveaway items such as pens, bags, and earphones to complete each survey.

Measures

All survey measures have been previously tested and validated. All measures were assessed through Web-based closed surveys, in the presence of a research assistant who was only available for technical assistance. We adhered to the Checklist for Reporting Results of Internet E-Surveys (CHERRIES; Multimedia Appendix 4) [56]. At baseline and follow-up, we measured intention to smoke [57,58] as the primary outcome. Intention to smoke, measured using the susceptibility to smoke scale, is the most potent and direct predictor of smoking initiation among adolescents [59,60]. At follow-up, we measured perceived interactivity [61], perceived entertainment [62,63], perceived presence [64], and emotional involvement [65,66]. We measured perceived credibility of message [67] for both conditions as a measure of manipulation check. To obtain a valid assessment of user experience and program effect, we included questions pertaining to potential confounders, such as environmental and social factors, and demographic characteristics that may contribute to smoking intention [68-73]: age, gender, ethnicity, average grade at school, the number of detentions at school, prior knowledge about the effects of smoking [74], number of friends who smoke, frequency of Internet use in hours per day, and skills in Internet use [75]. A detailed description of the main measures, measure derivation references, and Cronbach alpha values are reported in Table 2.

Table 2. Survey measures.

Measures	Description	Alpha[a]
Health outcome		
Intention to smoke	Items adapted from the susceptibility to smoke scale [58-60]: “Do you think in the future you might try a cigarette?”; “If one of your best friends were to offer you a cigarette, would you try it?”; and “Do you think you will try smoking some day in the next 5 years?” Items presented on a 5-point Likert scale with answer choices: 5=definitely yes, 4=probably yes, 3=possibly, 2=probably not, and 1=definitely not.	.80
User experience		
Perceived interactivity	Measured with 17 items from Coursaris and Sung [61], such as “I felt that I had control over my surfing experience of the website” and “The website seemed to be effective in getting my feedback.”	.94
Perceived entertainment	A scale adapted from the work of Cyr and colleagues [62] and Nysveen and colleagues [63], with two dimensions: enjoyment and fulfillment. Examples of statements: “I found my visit to this website entertaining” and “I found my visit to this website fun.”	.92
Media effect outcomes		
Perceived presence	Measured using 5-point Likert scale items such as “While using the website, I had a sense of being in the scenes” and “While using the website, I felt I was visiting the website’s world” [64].	.88
Emotional involvement	Two items: The first item belongs to the emotional involvement dimension of the transportation concept [65]: “I felt emotionally involved in the website.” The second item asked participants how much they felt emotionally involved in the website [66]. Answer choices ranged from 1=not at all to 10=extremely.	.55[b]
Covariates		
Prior knowledge	21 items tested knowledge about smoking consequences. Participants indicated if they believe such items are actual consequences of smoking by answering “yes” (coded 1 if correct), “no” (coded 1 if correct), or “I do not know” (always coded 0, as incorrect) [74].	-
Number of friends who smoke	One open-ended question: “How many of your friends smoke?”	-
Frequency of Internet use	One open-ended question: “How many hours per day do you spend on the internet?”	-
Number of school detentions	One open-ended question: “How many detentions or suspensions have you received at school?”	-
School grades	Total grade at school, based on grade point average. Answer choices were A, B, C, D, or F (coded 1, 2, 3, 4, and 5, respectively).	-
Perceived credibility	Two items such as “In your opinion, how believable was the information presented in ASPIRE?” [67].	

[a] Reliability coefficients with Cronbach alpha were calculated from posttest data, with the exception of measures with data collected at baseline only.

[b] Indicates Pearson correlation between 2 items, instead of Cronbach alpha.

Statistical Analysis

We conducted power analysis for sample size calculation. In order to conduct mixed-effects models using a target power of 0.95 and an effect size of 0.20 to predict intention to smoke [76] with an alpha value of .05, the estimated sample size is 80 participants (40 participants in each condition). Approximately 100 adolescents were needed to test these hypotheses, with an anticipated completion rate of 80% (80/100).

Statistical analyses were conducted using Stata version 14 (StataCorp LP). Analyses of variance (ANOVAs) and chi-square tests were conducted to capture any baseline differences between the 2 conditions with respect to covariates (eg, demographic characteristics, grades at school, number of detentions, and frequency of Internet use). Then, manipulation checks were conducted to ensure that both conditions are found to provide credible health information. Bonferroni adjustment was performed to guard against type I error in the repeated ANOVAs [77,78].

To test hypothesis 1, one-way ANOVA was conducted to assess whether using ASPIRE is related to perceived interactivity and perceived entertainment. For hypothesis 2, a repeated-measures mixed-effects model was conducted to test change in intention to smoke, in a 2 (condition) × 2 (time) design. ASPIRE effects on outcome trajectories over time were measured by the condition × time interaction term. To test hypothesis 3 on the effects of perceived interactivity and perceived entertainment, 3 repeated-measures mixed-effects models were conducted with intention to smoke as the outcome variable. Model 1 tested perceived interactivity alone, model 2 tested perceived entertainment alone, and model 3 included both as independent variables. For all models, multicollinearity was tested and the Huber-White sandwich estimator was used to correct all variance estimates for heteroskedasticity [79,80]. For each finding, standardized coefficients were computed with their respective *P* values.

Hypotheses 4 through 7 were first examined using mixed-effects models and multiple regression analyses that adjusted for effects of potential confounders, as covariates. These adjustments did not alter primary conclusions (data not shown). To confirm hypotheses 1 through 3 and test for hypotheses 4 through 7 under one model (Figure 1), path analysis was conducted. Path analysis has been regarded as a preferred method of mediation analysis as it allows the control for measurement error and uncovers information on the degree of fit of the entire model [81-86]. Model fit criteria were as follows: (1) a nonsignificant chi-square goodness-of-fit statistic, (2) a comparative fit index (CFI) of .90 or greater, (3) a root mean square error of approximation (RMSEA) less than or equal to .06, and (4) a standardized root mean square residual lower than .08 [87]. Two models were constructed with (1) perceived interactivity and (2) perceived entertainment as drivers of change. For each model, modification indices were examined to give insight into possible structural aspects of model misfit and update the model based on the indices.

Results

Participants

Table 3 presents respondents' sociodemographic characteristics. For the entire sample (N=101), the average age was 13.44 (SD 1.42) years, 43.56% (44/101) were female, and the majority of participants were black or African American (42/101, 41.58%) and Hispanic or Latino (44/101, 43.56%). Approximately 45% (43/101) reported having at least one friend who smokes, and the number of friends who smoke ranged between 1 and 23 (mean 1.96, SD 3.67). Also, 62.38% (63/101) reported using the Internet at least 2 hours per day, and the hours of use ranged between 30 minutes and 16 hours (mean 4.40, SD 3.70). The 2 groups differed with respect to participants' number of friends who smoke, $F_{1,94}$=5.74, *P*=.02. Subsequent analysis controlled for the difference.

Table 3. Characteristics of study participants.

Characteristics	ASPIRE[a] (n=50) n (%)	ASPIRE-control (n=51) n (%)	Total sample (N=101) n (%)	*P* value[b]
Age range, years				
12-13	26 (52)	38 (74.5)	64 (63.2)	.05
14-15	18 (36)	11 (21.5)	29 (28.7)	
16-17	6 (12)	2 (3.9)	8 (7.9)	
Gender				
Male	31 (60.7)	27 (52.9)	58 (56.8)	.42
Female	20 (39.2)	24 (47.0)	44 (43.5)	
Race or ethnicity				
Hispanic or African American	39 (78)	47 (92.1)	86 (85.1)	.05
Non-Hispanic, non–African American	11 (22)	4 (7.8)	15 (14.8)	
Educational level of mother				
High school or less	16 (32.6)	23 (45.1)	39 (39.0)	.43
College or more	33 (67.3)	28 (54.9)	61 (61.0)	
Educational level of father				
High school or less	23 (46.9)	29 (61.7)	52 (54.1)	.27
College or more	26 (53.0)	18 (38.3)	44 (45.8)	
Educational level of legal guardian				
High school or less	9 (24.3)	12 (46.1)	21 (33.3)	.16
College or more	28 (75.6)	14 (53.8)	42 (66.6)	
Average school grades				
A	37 (72.5)	31 (60.7)	68 (66.6)	.49
B	12 (23.5)	17 (33.3)	29 (28.4)	
C	2 (3.9)	2 (3.9)	4 (3.9)	
D	0 (0.0)	1 (1.9)	1 (0.9)	
Number of school detentions, mean (SD)	1.52 (3.95)	1.69 (2.65)	1.61 (3.34)	.80
Number of friends who smoke, mean (SD)	1.06 (1.98)	2.82 (4.62)	1.96 (3.67)	.02
Prior knowledge of smoking effects, mean (SD)	14.04 (3.80)	12.94 (4.03)	13.49 (3.94)	.16
Prior intention to smoke, mean (SD)	1.43 (0.65)	1.56 (0.72)	1.50 (0.68)	.36
Frequency of Internet use, mean (SD)	3.77 (3.48)	5.01 (3.83)	4.40 (3.70)	.10

[a] ASPIRE: A Smoking Prevention Interactive Experience.

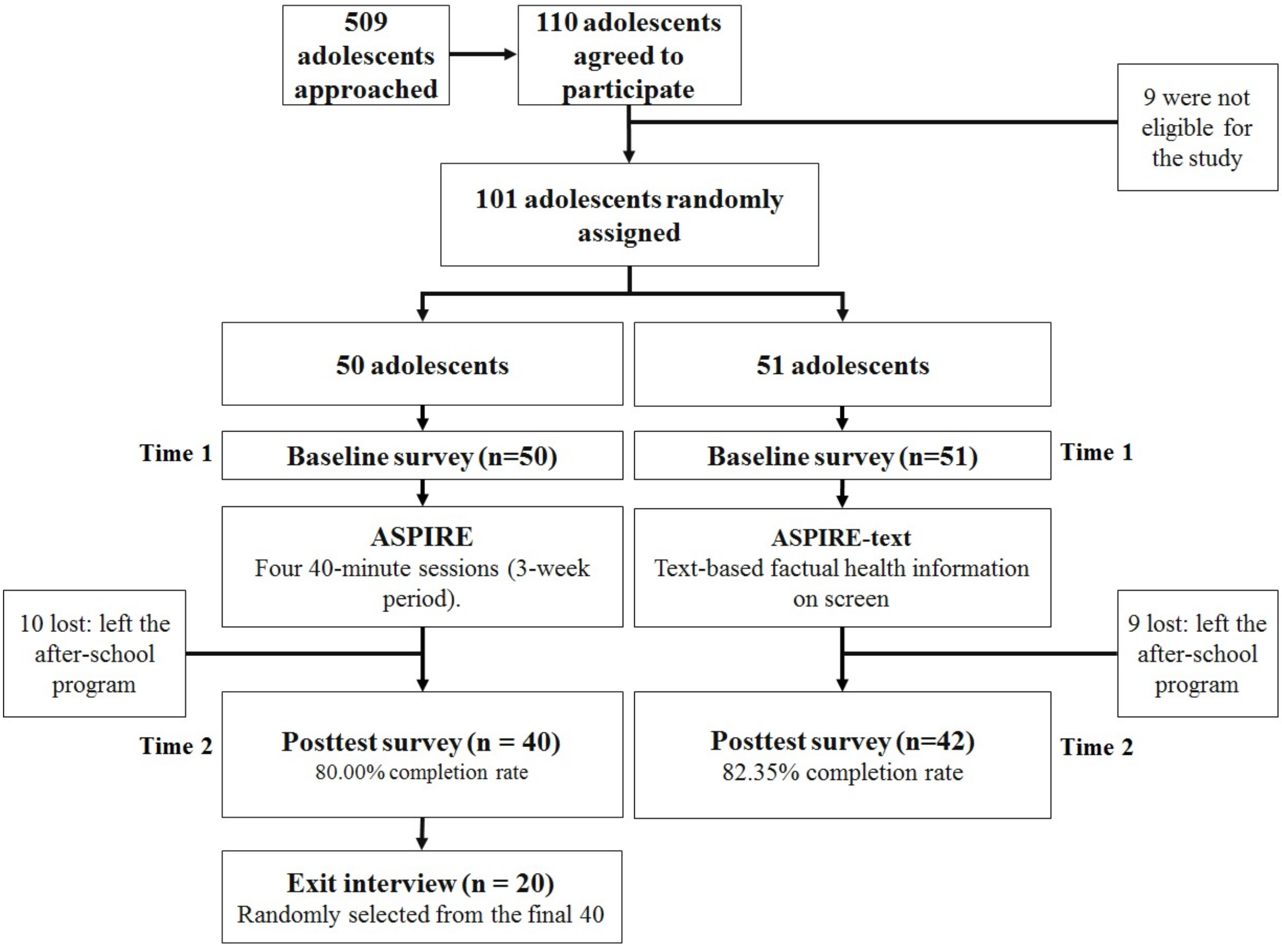
[b] Significance testing with chi-square test for the categorical variables (ie, age, gender, race or ethnicity, educational level, and school grades) and analysis of variance for the continuous variables. Missing values are not presented in this table.

Attrition

A total of 110 adolescents agreed to participate. We excluded 9 adolescents who did not meet the adolescent age criterion (ages 12 through 18 years). A total of 101 participants took the baseline survey and were randomized to 1 of the 2 conditions. All 101 participants went through ASPIRE and ASPIRE-control as prescribed and completed all sessions. Then, 81.20% (82/101) continued until follow-up (81.2% completion rate; Figure 2). Participants who did not continue were adolescents who have left their after-school program.

There was no significant difference between participants who did and those who did not continue to follow-up with respect to baseline intention to smoke ($F_{1,99}$=0.03, *P*=.87), prior knowledge about smoking outcomes ($F_{1,100}$=0.13, *P*=.72), number of friends who smoke ($F_{1,94}$=0.09, *P*=.76), number of detentions at school ($F_{1,95}$=0.70, *P*=.40), age ($F_{1,99}$=1.15, *P*=.29), gender (χ^2_1=0.03, *P*=.85), ethnicity (χ^2_1=0.5 *P*=.50), prior frequency of Internet use ($F_{1,96}$=2.72, *P*=.10), and skills in Internet use ($F_{1,100}$=1.76, *P*=.19).

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. The exit interview data are not included in this paper. ASPIRE: A Smoking Prevention Interactive Experience.



Manipulation Checks

Perceived Credibility of Health Information

We checked to make sure that the conditions present credible health messages. As expected, there was no significant difference between the conditions with regard to perceived credibility of message content, $F_{1,56}$=3.67, P=.06. Perceived credibility scores for both groups were found to be significantly greater than 3 on a 5-point Likert scale (ASPIRE: t_{40}=4.38, P<.001; control: t_{42}=8.48, P<.001). This confirms that both interventions were perceived to be credible sources of information related to smoking.

Checking for Baseline Differences

There was a marginal difference between the groups with respect to age, ethnicity, and number of friends who smoke (Table 3). When controlled, these differences did not affect the results of this study. The 2 groups did not differ with respect to the number of school detentions, prior knowledge about the effects of smoking, or baseline level of intention to smoke (Table 3).

Checking for the Digital Divide

The results indicate that the 2 groups did not differ with respect to the frequency of Internet use in hours per day (Table 3). There were no differences between ethnicities ($F_{1,93}$=0.12, P=.73), genders ($F_{1,96}$=0.60, P=.44), those who have and those who have not had detentions ($F_{1,93}$=1.32, P=.25), or those who have and those who have not had friends who smoke ($F_{1,91}$=3.30, P=.07), with respect to the frequency of Internet use.

Checking for Confounders

To check for potential demographic confounders, we determined whether intervention effects varied by demographic characteristics, using moderation analysis with mixed-effects models. Overall, the results failed to identify differential effects as a function of age (P=.11), ethnicity (P=.43), number of detentions (P=.55), frequency of Internet use (P=.39), knowledge about smoking effects (P=.36), or number of friends who smoke (P=.76). Being male (P=.01) and having higher grades (P=.02) moderated the effect of ASPIRE on intention to smoke.

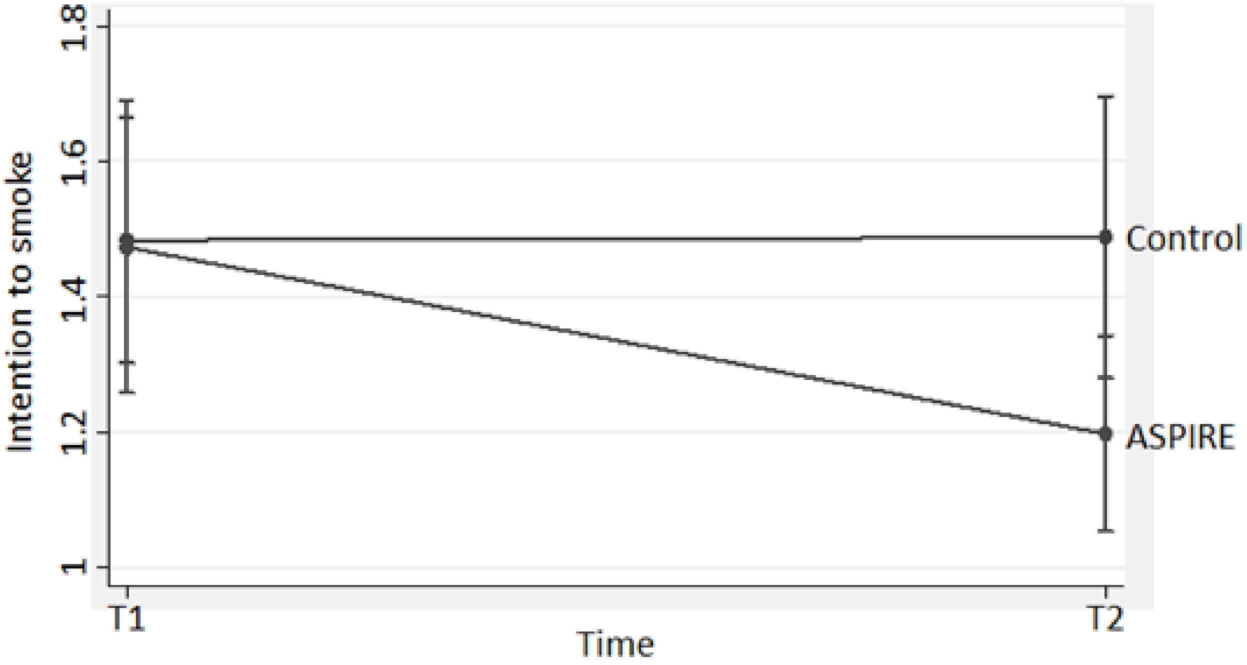
From Design Manipulation to User Experience

To test hypothesis 1, we checked to see if the manipulation of interactivity and entertainment has led to an experience of interactivity and entertainment among adolescent users. There was a significant difference between the ASPIRE group and the ASPIRE-control group with respect to perceived interactivity ($F_{1,82}$=11.66, P=.001) and perceived entertainment ($F_{1,81}$=16.40, P<.001).

Intention to Smoke

A mixed-effects model predicting intention to smoke and controlling for confounders showed support for hypothesis 2 (Figure 3). Adolescents in the ASPIRE group were significantly more likely to show a decrease in their intention to smoke over time compared with participants in the ASPIRE-control group (group × time interaction effect; beta=−0.18, *P*=.008). The ASPIRE group showed a significant decrease in intention to smoke (slope=−0.28, *P*=.008), while the ASPIRE-control group showed no significant change from baseline to follow-up (slope<0.001, *P*=.99).

Figure 3. Change in intention to smoke over time for ASPIRE (A Smoking Prevention Interactive Experience) and ASPIRE-control.



Experience of Interactivity and Entertainment

There was a significant correlation between perceived interactivity and perceived entertainment (*r*=.68, *P*<.001). This correlation is expected, considering that interactivity and entertainment in ASPIRE overlap through user activities. Supporting hypothesis 3, mixed-effects models of Table 4 showed that perceived interactivity and perceived entertainment independently worked as significant predictors of a decrease in intention to smoke (model 1: beta$_{\text{interactivity}}$=−0.27, *P*=.004 and model 2: beta$_{\text{entertainment}}$=−0.20, *P*=.04). When both factors are included in the model (model 3), only perceived interactivity significantly predicted a decrease in intention to smoke (model 3: beta$_{\text{interactivity}}$=−0.23, *P*=.04; beta$_{\text{entertainment}}$=−0.06, *P*=.61). In all 3 models, the group-by-time interaction effect maintained significance (Table 4).

Table 4. Repeated-measures mixed-effects models with perceived interactivity and perceived entertainment predicting intention to smoke.

Variables	Intention to smoke[a]					
	Model 1		Model 2		Model 3	
	beta (SE)[b]	*P* value	beta (SE)	*P* value	beta (SE)	*P* value
Perceived interactivity[c]	−0.27 (0.01)	.004	-	.039	-0.23 (0.01)	
Perceived entertainment[c]	-		−0.20 (0.01)	.038	−0.06 (0.01)	.60
Condition	0.11 (0.17)	.355	0.10 (0.18)	.462	0.16 (0.18)	.37
Time	0.01 (0.06)	.797	0.01 (0.06)	.856	0.01 (0.06)	.85
Condition × time[c]	−0.22 (0.11)	.007	−0.19 (0.11)	.007	−0.19 (0.11)	.007
Intercept	3.05 (0.81)	<.001	2.92 (0.83)	<.001	3.16 (0.89)	<.001
Wald chi-square	34.32	<.001	35.04	<.001	36.68	<.001

[a] Indicates dependent outcome variable.

[b] Indicates standardized values followed by standard error.

[c] Indicates variables of interest. The models control for covariates (age, gender, prior knowledge, school grades, school detentions, and number of friends who smoke), with no significant relationship between such covariates and intention to smoke.

Path Models

Controlling for baseline group differences and confounders, path model results remained the same. With perceived entertainment (path model 1, Figure 4), path analysis showed that adolescents in the ASPIRE condition were more likely to experience entertainment than adolescents in the control group (beta=0.27, *P*<.001). The relationship between ASPIRE use and emotional involvement was lost when perceived entertainment was added to the model (beta=0.06, *P*=.40). Perceived entertainment fully mediated this relationship, indicating support for hypothesis 4. In model 1, perceived entertainment was significantly related to emotional involvement (beta=0.50, *P*<.001). This relationship remained when perceived presence was added to the model (beta=0.62, *P*<.001). Also, perceived presence was not related to emotional involvement (beta=0.19, *P*=.05). Unsupportive of hypothesis 5, perceived presence did not mediate the relationship between perceived entertainment and emotional involvement. Following this path in model 1, emotional involvement was related to lower intention to smoke (beta=−0.16, *P*=.01; hypothesis 7). Overall, model 1 presented acceptable fit when predicting intention to smoke (χ^2_3=5.3, *P*=.142, CFI=.9, RMSEA=.07, 90% CI .00-.17). The model explained 9.7% of the variance in intention to smoke (coefficient of determination, CD=0.097). There were no modification indices needed to improve this model.

With perceived interactivity (path model 2, Figure 4), path analysis showed that adolescents in the ASPIRE condition were more likely to experience interactivity than adolescents in the control group (beta=0.24, *P*=.003). The relationship between ASPIRE use and emotional involvement was lost when perceived interactivity was added to the model (beta=0.13, *P*=.10), indicating full mediation that supports hypothesis 4. The relationship between perceived interactivity and emotional involvement was weakened when perceived presence was added to the model as a mediator (beta=0.20, *P*=.01). Testing for hypothesis 6, perceived presence partially mediated the relationship between the experience of interactivity and emotional involvement. Following this mediation, emotional involvement was related to lower intention to smoke (beta=−0.16, *P*=.01; hypothesis 7). Overall, the model offered poor fit predicting intention to smoke (χ^2_3=9.8, *P*=.02, CFI=.9, RMSEA=.123, 90% CI .04-.21). The model explained 9.90% of the variance in intention to smoke (CD=0.099). Modification indices called for the addition of a causal path from perceived interactivity to intention to smoke. As a result, model 3 was constructed with this addition.

Path model 3 of Figure 4 indicated a significant direct relationship between perceived interactivity and intention to smoke (beta=−0.25, *P*<.001). The relationship between emotional involvement and intention to smoke lost significance as a result of this addition (beta=−0.05, *P*=.50). All other paths kept significance. Model 3 indicated acceptable fit when predicting intention to smoke (χ^2_2=1.8, *P*=.41, CFI=1.0, RMSEA=.00, 90% CI .00-.16). The model explained 9.0% of the variance in intention to smoke (CD=0.090). There were no modification indices needed to improve this model.

Figure 4. Path models indicating the path from ASPIRE (A Smoking Prevention Interactive Experience) use to intention to smoke. Note. By controlling for the effect of confounders and demographic group differences at baseline, the results remained the same. CFI: comparative fit index; RMSEA: root mean square error of approximation * *P*<0.05; ** *P*<0.01; *** *P*<0.001; † *P*<0.1.

Discussion

Conclusions

This post hoc study of a randomized controlled trial examined the process by which adolescents' experience with a Web-based smoking prevention intervention leads to a health outcome. This study was the first step to better understand the underlying mechanism of eHealth effects, with a user-centered approach. The study identified salient variables in this mechanism from the perspective of the users' experience and then postulated and empirically tested a model that can help explain how this mechanism takes place cognitively and emotionally, to reach the intended health outcome.

In particular, while ASPIRE has previously shown success in delaying smoking initiation [19], this study showed how adolescents' experience with ASPIRE contributes to expected outcomes such as decreased intention to smoke. The findings expand previous efforts to explain the effects of Web-based interventions. They (1) go beyond focus on the health outcomes and (2) explicate associations that allow emotional involvement to be a driver of smoking prevention.

The results demonstrate the advantages of a more engaging user experience perceived as interactive and entertaining over the mere exposure to health information in conventional interventions. The 2 critical elements of the user experience manipulated in this study are interactivity and entertainment, as ASPIRE features activities with real-time feedback and entertaining videos. As expected, such manipulation predicted a perception of interactivity and entertainment. The more adolescents found ASPIRE to be interactive and entertaining, the more they were likely to show a decrease in their intention to smoke. A closer look at these findings shows that perceived interactivity in ASPIRE has a stronger relationship with the decrease in intention to smoke than perceived entertainment. This may not necessarily be due to the lack of entertainment in ASPIRE. Instead, this finding may be due to adolescents' tendency to prefer first-hand involvement in activities and experiential learning instead of exposure to entertaining videos. As a result, Web-based smoking prevention programs that use entertainment need to concentrate their efforts on interactive elements that can boost the effect of entertainment on adolescents' intention to smoke.

The results also suggest that adolescents tend to transition from the experience of entertainment and interactivity to a psychological state of reduced intention to smoke, passing through emotional involvement. Adolescents transitioned from the experience of media elements (interactivity and entertainment) to perceived presence, emotional involvement, and ultimately lower intention to smoke. This model suggests that the elements of Web-based interventions for smoking prevention (eg, interactivity and entertainment) can have specific psychological effects beyond those explained by cognitive theories (eg, the social learning theory and the health belief model). In particular, concepts such as perceived presence and emotional involvement contribute to predicting tobacco control outcomes.

Furthermore, perceived interactivity and perceived entertainment took separate paths to reach emotional involvement in ASPIRE. Perceived interactivity did not directly correlate with a state of emotional involvement in ASPIRE. Instead, perceived presence allowed for an indirect relationship between perceived interactivity and emotional involvement. The more adolescents found ASPIRE to be interactive, the more they felt present in the ASPIRE environment. Then, perceived presence drove emotional involvement in ASPIRE. On the other hand, perceived entertainment exhibited a direct relationship with emotional involvement, consistent with previous research that presents entertainment as a driver of emotions through drama [88], humor [89], and horror [90].

The findings support the notion that emotional involvement is important if we are to impact adolescent intentions. Emotional involvement seems to play a role in ASPIRE by bridging the gap between experience and smoking intentions. In addition to user activities, the videos in ASPIRE can have a strong emotional influence. Such videos portray dramatic stories through testimonials from smokers facing the effects of smoking, humoristic stories about social situations and smoking outcomes, and fear appeal through the depiction of oral and maxillofacial cancer as a result of smoking. Regardless of the user experience, entertainment and interactivity seem to elicit emotions that predict a decrease in intention to smoke. This supports previous work suggesting that messages with an emotional tone have an impact on youth smoking behavior [91,92]. Future research can expand this work to explore the psychological processes that move adolescents from emotional involvement to lower intention to smoke. One unexpected finding is the enhancement of path model 2 through the addition of a direct relationship between perceived interactivity and intention to smoke. Added in path model 3, this relationship rendered the effect of emotional involvement on intention to smoke nonsignificant. This result corroborates the important role of interactivity in driving health outcomes.

Limitations

Some limitations for this study must be noted. Participants in both conditions were asked to sit in front of laptops and follow a stringent regimen of ASPIRE usage. During each session, they had to keep an unchanged sitting position without peer-to-peer interaction. While this procedure did not allow adolescents to behave as in a natural environment, it created a uniform Web-based experience that controls for any contamination of results that may be due to distractions during intervention use.

Although the study design involved a manipulation of ASPIRE that mechanically removed entertainment and interactivity from the intervention to create the control condition, the study design did not separate interactivity from entertainment. ASPIRE was designed in a Flash Player format, which is costly to manipulate. In addition, several activities in ASPIRE involved an amalgam between interactivity and entertainment. As a result, the separation between these features in ASPIRE is not sound because it can disrupt the overall ASPIRE experience. In the future and outside the context of ASPIRE, we plan to conduct a study that physically separates interactivity from entertainment and compares the 2 features with each other.

The findings of the transitional model must be interpreted with caution, considering that data analysis is conducted through regression and causation cannot be inferred (eg, change over time in emotional involvement). Nevertheless, path analysis supports the presence of transition that can be confirmed in the future, through a causation model.

Although the study predicted intention to smoke, it did not consider long-term opportunities for protective behaviors or the measurement of actual smoking initiation. While smoking initiation was already examined during the main randomized controlled trial for ASPIRE [19], this study was meant to study the role of adolescents' experience with ASPIRE in driving smoking-related outcomes. Also, intention to smoke is the closest, most potent predictor of smoking initiation [59,60].

Finally, it must be noted that the opportunity to provide a deep and comprehensive analysis is limited by the relatively small sample size. Future work may consider examining how the ASPIRE experience can prevent long-term smoking initiation with a larger sample of adolescents.

Implications

Several implications for future research and practice can be envisioned. The results of this study indicate that researchers can work to enhance the experience of interactivity and entertainment in order to better reach outcomes related to nicotine and tobacco control. First, the study of intervention experience can ultimately inform new ways to improve content of Web-based smoking prevention interventions. Through a collaboration with health communication scientists, public health researchers can work to maximize emotional involvement through the elements of interactivity and entertainment. Second, the results encourage further investigation of entertainment in order to find ways to improve its impact. Knowing that interactivity plays an important role in reducing intention to smoke, intervention designers can make use of entertaining features in an interactive environment. For instance, ASPIRE researchers may introduce game-based activities into their interventions. By incorporating game elements into purely interactive activities, adolescents may be transported into a playful environment that can increase their emotional involvement [93,94]. In the future, we plan to examine further how adolescents go from a state of emotional involvement to a decrease in intention to smoke. We plan to also examine how negative and positive emotions and discrete emotions (eg, happiness, sadness, surprise, and fear) may differ in their effects on intention to smoke. We also plan to study further the threshold for entertainment and interactivity required for the program to effectively influence intention to smoke. Finally, future endeavors may involve the application of the user experience model in the context of other health behaviors in order to test its generalizability.

Acknowledgments

We thank the advising committee from the University at Buffalo, the State University of New York (Buffalo, New York) for their contributions to the success of this trial, including Mark Frank, PhD, and Lance Rintamaki, PhD. We also thank the University of Texas MD Anderson Cancer Center (creators of ASPIRE, Houston, Texas). Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R25CA057730 (Principal Investigator: Shine Chang, PhD) and by the Cancer Center Support Grant CA016672 (Principal Investigator: Ronald DePinho, MD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

GEK is responsible for the design of the study; AVP and HW provided guidance on the design of the study; GEK was responsible for the data collection, acquisition, and analysis; NM and KSC participated in the data collection and implementation of the study procedure at each research site; GEK, AVP, HW, NM, KSC, and RS contributed to the conceptualization and design of the paper; GEK drafted the paper; HW, AVP, KSC, NM, and RS critically revised the paper. All authors read and approved the final version. GEK had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the ASPIRE (A Smoking Prevention Interactive Experience) and ASPIRE-control webpages.

[PNG File, 149KB - jmir_v19i2e44_app1.png]

Multimedia Appendix 2

Layout and design of A Smoking Prevention Interactive Experience (ASPIRE).

[PNG File, 1MB - jmir_v19i2e44_app2.png]

Multimedia Appendix 3

A video of the ASPIRE (A Smoking Prevention Interactive Experience) website presenting an activity in action.

[MP4 File (MP4 Video), 6MB - jmir_v19i2e44_app3.mp4]

Multimedia Appendix 4

Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

[PDF File (Adobe PDF File), 35KB - jmir_v19i2e44_app4.pdf]

Multimedia Appendix 5

CONSORT eHealth V1.6.1 checklist.

[PDF File (Adobe PDF File), 534KB - jmir_v19i2e44_app5.pdf]

References

1. US Department of Health and Human Services. Surgeongeneral.gov. 2014. The health consequences of smoking-50 years of progress: a report of the Surgeon General URL: https://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf
2. National Institute on Drug Abuse. Drugabuse.gov. 2015. Monitoring the Future Study: Trends in Prevalence of Various Drugs URL: https://www.drugabuse.gov/trends-statistics/monitoring-future/monitoring-future-study-trends-in-prevalence-various-drugs [accessed 2016-12-05] [WebCite Cache ID 6mXC5LHn3]
3. Miech R, Johnson L, O?Malley P. Monitoringthefuture.: The University of Michigan Institute for Social Research; 2015. Monitoring the Future national survey results on drug use, 1975-2014 Volume I: Secondary school students URL: http://www.monitoringthefuture.org/pubs/monographs/mtf-vol1_2014.pdf [accessed 2016-12-05] [WebCite Cache ID 6mXDXcx5T]
4. Singh T, Arrazola RA, Corey CG, Husten CG, Neff LJ, Homa DM, et al. Tobacco use among middle and high school students--United States, 2011-2015. MMWR Morb Mortal Wkly Rep 2016 Apr 15;65(14):361-367 [FREE Full text] [doi: 10.15585/mmwr.mm6514a1] [Medline: 27077789]
5. Marshall JR, Lotfipour S, Chakravarthy B. Growing trend of alternative tobacco use among the nation's youth: a new generation of addicts. West J Emerg Med 2016 Mar;17(2):139-142 [FREE Full text] [doi: 10.5811/westjem.2016.1.29383] [Medline: 26973738]
6. National Cancer Institute. Progressreport.cancer. 2012. Youth Tobacco Use URL: https://www.progressreport.cancer.gov/prevention/youth_smoking [accessed 2016-12-05] [WebCite Cache ID 6mXNO0hsg]
7. Eng T. The ehealth landscape: a terrain map of emerging information and communication technologies. Princeton, N.J: Robert Wood Johnson Foundation; 2001.
8. US Department of Health and Human Services. Health. Washington, DC: Office of Disease Prevention and Health Promotion; 2006. Expanding the reach of consumer e-health tools URL: https://health.gov/communication/ehealth/ehealthtools/default.htm [accessed 2016-12-05] [WebCite Cache ID 6mXOC8hNY]
9. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
10. Wiehe SE, Garrison MM, Christakis DA, Ebel BE, Rivara FP. A systematic review of school-based smoking prevention trials with long-term follow-up. J Adolesc Health 2005 Mar;36(3):162-169. [doi: 10.1016/j.jadohealth.2004.12.003] [Medline: 15737770]
11. Lustria Mia Liza A, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
12. Sussman S, Sun P, Dent CW. A meta-analysis of teen cigarette smoking cessation. Health Psychol 2006 Sep;25(5):549-557. [doi: 10.1037/0278-6133.25.5.549] [Medline: 17014271]
13. Cremers H, Mercken L, Candel M, de Vries H, Oenema A. A web-based, computer-tailored smoking prevention program to prevent children from starting to smoke after transferring to secondary school: randomized controlled trial. J Med Internet Res 2015 Mar 09;17(3):e59 [FREE Full text] [doi: 10.2196/jmir.3794] [Medline: 25759248]
14. de Josselin de Jong S, Candel M, Segaar D, Cremers H, de Vries H. Efficacy of a Web-based computer-tailored smoking prevention intervention for Dutch adolescents: randomized controlled trial. J Med Internet Res 2014 Mar 21;16(3):e82 [FREE Full text] [doi: 10.2196/jmir.2469] [Medline: 24657434]
15. Tingle LR, DeSimone M, Covington B. A meta-evaluation of 11 school-based smoking prevention programs. J Sch Health 2003 Feb;73(2):64-67. [Medline: 12643021]
16. Valente T. Evaluating health promotion programs. Oxford: Oxford University Press; 2002.
17. Shegog R, Prokhorov A, Kelder S. Project ASPIRE: Intermediate results from a multimedia smoking preventioncessation program for adolescents. 2004 Presented at: The 132nd Annual Meeting of the American Public Health Association; 11/06/2004; Washington, DC.
18. Prokhorov AV, Yost T, Mullin-Jones M, de Moor C, Ford KH, Marani S, et al. “Look at your health”: outcomes associated with a computer-assisted smoking cessation counseling intervention for community college students. Addict Behav 2008 Jun;33(6):757-771 [FREE Full text] [doi: 10.1016/j.addbeh.2007.12.005] [Medline: 18280668]
19. Prokhorov AV, Kelder SH, Shegog R, Murray N, Peters R, Agurcia-Parker C, et al. Impact of A Smoking Prevention Interactive Experience (ASPIRE), an interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high-school students. Nicotine Tob Res 2008 Sep;10(9):1477-1485. [doi: 10.1080/14622200802323183] [Medline: 19023839]

20. Prokhorov AV, Kelder SH, Shegog R, Conroy JL, Murray N, Peters R, et al. Project ASPIRE: An interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high school students. Subst Use Misuse 2010 May;45(6):983-1006. [doi: 10.3109/10826080903038050] [Medline: 20397881]
21. Sundar S, Xu Q, Bellur S. Designing interactivity in media interfaces: A communications perspective. In: Proceedings of the SIGCHI 28th International Conference on Human Factors in Computing Systems. Designing interactivity in media interfaces; 2010 Presented at: The SIGCHI 28th International Conference on Human Factors in Computing Systems; 2010; Atlanta, GA. [doi: 10.1145/1753326.1753666]
22. Sayre S, King C. Entertainment and society: influences, impacts, and innovations. New York: Routledge; 2010.
23. Sayre S, King C. Entertainment & society: audiences, trends, and impacts. Thousand Oaks, Calif: Sage Publications; 2003.
24. Vogel H. Entertainment industry economics: a guide for financial analysis. Cambridge: Cambridge University Press; 2001.
25. Vorderer P, Steen F, Chan E. Motivation. In: Bryant J, Vorderer P, editors. Psychology of entertainment. Mahwah, NJ: Lawrence Erlbaum Associates; 2006.
26. Bucy EP. Interactivity in society: locating an elusive concept. Inform Soc 2004 Nov;20(5):373-383. [doi: 10.1080/01972240490508063]
27. Kiousis S. Interactivity: a concept explication. New Media Soc 2002 Sep 01;4(3):355-383. [doi: 10.1177/146144480200400303]
28. McMillan SJ, Hwang J. Measures of perceived interactivity: an exploration of the role of direction of communication, user control, and time in shaping perceptions of interactivity. J Advertising 2013 May 31;31(3):29-42. [doi: 10.1080/00913367.2002.10673674]
29. Laurel B. Computers as Theater. Boston: Addison-Wesley; 1991.
30. Bosshart L, Macconi I. Media entertainment. Commun Res Trends 1998;18:3-6 [FREE Full text]
31. Bae H, Kang S. The roles of interactivity and involvement in e-Health. Hum Res 2009;56:227-266 [FREE Full text]
32. Joinson A, McKenna K, Postmes T, Reips U. Assessing interactivity in computer-mediated research. In: Joinson A, McKenna T, Reips U, editors. The Oxford handbook of Internet psychology. Oxford: Oxford University Press; 2007.
33. Noar SM, Harrington NG, editors. eHealth Applications: Promising Strategies for Behavior Change (Routledge Communication Series). New York, NY: Routledge; 2012.
34. Heijden H. Research Memorandum. 2000. Using the technology acceptance model to predict website usagextensions and empirical test URL: http://degree.ubvu.vu.nl/repec/vua/wpaper/pdf/20000025.pdf [WebCite Cache ID 6mYW8n0D4]
35. Kim J, Forsythe S. Sensory enabling technology acceptance model (SE-TAM): A multiple-group structural model comparison. Psychol Mark 2008 Sep;25(9):901-922. [doi: 10.1002/mar.20245]
36. Yang B, Kim Y, Yoo C. The integrated mobile advertising model: the effects of technology- and emotion-based evaluations. J Bus Res 2013 Sep;66(9):1345-1352. [doi: 10.1016/j.jbusres.2012.02.035]
37. Bryant J, Vorderer P. Psychology of entertainment. Mahway, NJ: Lawrence Erlbaum; 2006.
38. Kolb A, Kolb D. Experiential learning theory. In: Seel NM, editor. Encyclopedia of the Sciences of Learning. New York, NY: Springer; 2012:1215-1219.
39. Slater MD, Rouner D. Entertainment-education and elaboration likelihood: understanding the processing of narrative persuasion. Commun Theory 2002 May;12(2):173-191. [doi: 10.1111/j.1468-2885.2002.tb00265.x]
40. Mutlu B, Forlizzi J, Nourbakhsh I, Hodgins J. The use of abstractionmotion in the design of social interfaces. In: Proceedings of the 6th conference on Designing Interactive systems. New York: ACM; 2006 Presented at: The 6th conference on Designing Interactive systems; 06/26/2006; University Park, PA. [doi: 10.1145/1142405.1142444]
41. Lim Y, Lee S, Lee K. Interactivity attributes: a new way of thinking and describing interactivity. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 Presented at: The SIGCHI Conference on Human Factors in Computing Systems; 2009; Boston, MA p. 04-09. [doi: 10.1145/1518701.1518719]
42. Sundar SS, Wagner CB. The world wide wait: exploring physiological and behavioral effects of download speed. Media Psychol 2002 May;4(2):173-206. [doi: 10.1207/S1532785XMEP0402_04]
43. Lang A, Borse J, Wise K, David P. Captured by the World Wide Web: orienting to structural and content features of computer-presented information. Commun Res 2002 Jun 01;29(3):215-245. [doi: 10.1177/0093650202029003001]
44. Wise K, Reeves B. The effect of user control on the cognitive and emotional processing of pictures. Media Psychol 2007 May 15;9(3):549-566. [doi: 10.1080/15213260701283186]
45. Tammelin M. Hrast.pef. 1998. From telepresence to social presence: The role of presence in a network-based learning environment URL: http://hrast.pef.uni-lj.si/~joze/podiplomci/FRI/Tammelin.pdf [accessed 2017-02-13] [WebCite Cache ID 6oF5ZunDf]
46. Perse EM. Audience selectivity and involvement in the newer media environment. Commun Res 1990 Oct 01;17(5):675-697. [doi: 10.1177/009365090017005005]
47. Signorielli N, Morgan M. Cultivation and involvement with local television news. In: Signorielli N, Morgan M, editors. Cultivation analysis: new directions in media effects research. Newbury Park, CA: Sage Publications; 1990.
48. Perse EM. Involvement with local television news cognitive and emotional dimensions. Human Comm Res 1990 Jun;16(4):556-581. [doi: 10.1111/j.1468-2958.1990.tb00222.x]
49. Step M. An emotional appraisal model of media involvement, uses, and effects. Ann Arbor, MI: UMI; 1998.

50. Riva G, Mantovani F, Capideville CS, Preziosa A, Morganti F, Villani D, et al. Affective interactions using virtual reality: the link between presence and emotions. Cyberpsychol Behav 2007 Feb;10(1):45-56. [doi: 10.1089/cpb.2006.9993] [Medline: 17305448]
51. Green MC, Clark JL. Transportation into narrative worlds: implications for entertainment media influences on tobacco use. Addiction 2013 Mar;108(3):477-484. [doi: 10.1111/j.1360-0443.2012.04088.x] [Medline: 22994374]
52. Eysenbach G, Group CE. 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]
53. Eysenbach G. CONSORT-EHEALTH: implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [Medline: 23920638]
54. MD Anderson Cancer Center. MDAnderson.org. 2016. ASPIRE URL: https://www.mdanderson.org/about-md-anderson/community-services/aspire.html [accessed 2016-12-06] [WebCite Cache ID 6mYdlrpoO]
55. Steuer J. Defining virtual reality: dimensions determining telepresence. J Commun 2006;42:73-93. [doi: 10.1111/j.1460-2466.1992.tb00812.x]
56. Eysenbach G. Improving the quality of Web surveys: the checklist for reporting results of Internet e-surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
57. Straub DM, Hills NK, Thompson PJ, Moscicki AB. Effects of pro- and anti-tobacco advertising on nonsmoking adolescents' intentions to smoke. J Adolesc Health 2003 Jan;32(1):36-43. [Medline: 12507799]
58. Pierce JP, Choi WS, Gilpin EA, Farkas AJ, Merritt RK. Validation of susceptibility as a predictor of which adolescents take up smoking in the United States. Health Psychol 1996 Sep;15(5):355-361. [Medline: 8891714]
59. Choi WS, Gilpin EA, Farkas AJ, Pierce JP. Determining the probability of future smoking among adolescents. Addiction 2001 Feb;96(2):313-323. [doi: 10.1080/09652140020021053] [Medline: 11182877]
60. Jackson C. Cognitive susceptibility to smoking and initiation of smoking during childhood: a longitudinal study. Prev Med 1998;27(1):129-134. [doi: 10.1006/pmed.1997.0255] [Medline: 9465363]
61. Coursaris CK, Sung J. Antecedents and consequents of a mobile website's interactivity. New Media Soc 2012 Apr 26;14(7):1128-1146. [doi: 10.1177/1461444812439552]
62. Cyr D, Head M, Ivanov A. Perceived interactivity leading to e-loyalty: development of a model for cognitive–affective user responses. Int J Hum-Comput St 2009 Oct;67(10):850-869. [doi: 10.1016/j.ijhcs.2009.07.004]
63. Nysveen H, Pedersen PE, Thorbjørnsen H. Intentions to use mobile services: antecedents and cross-service comparisons. J Acad Market Sci 2005 Jul 01;33(3):330-346. [doi: 10.1177/0092070305276149]
64. Slater MD, Usoh M, Steed A. Depth of Presence in Virtual Environments. Presence Teleop Virt 1994 Jan;3(2):130-144. [doi: 10.1162/pres.1994.3.2.130]
65. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. J Pers Soc Psychol 2000 Nov;79(5):701-721. [Medline: 11079236]
66. Hess U, Kleck RE. Differentiating emotion elicited and deliberate emotional facial expressions. Eur J Soc Psychol 1990 Sep;20(5):369-385. [doi: 10.1002/ejsp.2420200502]
67. Eastin M. Credibility assessments of online health information: the effects of source expertise and knowledge of content. J Comput-Mediat Comm 2001;6:-. [doi: 10.1111/j.1083-6101.2001.tb00126.x]
68. Oenema A, Brug J, Lechner L. Web-based tailored nutrition education: results of a randomized controlled trial. Health Educ Res 2001 Dec;16(6):647-660 [FREE Full text] [Medline: 11780705]
69. Grandpre J, Alvaro EM, Burgoon M, Miller CH, Hall JR. Adolescent reactance and anti-smoking campaigns: a theoretical approach. Health Commun 2003;15(3):349-366. [doi: 10.1207/S15327027HC1503_6] [Medline: 12788679]
70. Zhao X, Cai X. The role of ambivalence in college nonsmokers' information seeking and information processing. Commun Res 2008 Apr 14;35(3):298-318. [doi: 10.1177/0093650208315959]
71. Van Deursen A, Van Dijk J. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
72. National Cancer Institute. Cancercontrol.cancer. Monograph 14: Changing adolescent smoking prevalence URL: http://cancercontrol.cancer.gov/brp/tcrb/monographs/14/ [WebCite Cache ID 6mYljqRQa]
73. Cokkinides VE, Halpern MT, Barbeau EM, Ward E, Thun MJ. Racial and ethnic disparities in smoking-cessation interventions: analysis of the 2005 National Health Interview Survey. Am J Prev Med 2008 May;34(5):404-412. [doi: 10.1016/j.amepre.2008.02.003] [Medline: 18407007]
74. Anjum Q, Ahmed F, Ashfaq T. Knowledge, attitude and perception of water pipe smoking (Shisha) among adolescents aged 14-19 years. J Pak Med Assoc 2008 Jun;58(6):312-317 [FREE Full text] [Medline: 18988390]
75. Teo TS, Lim VK, Lai RY. Intrinsic and extrinsic motivation in Internet usage. Omega 1999 Feb;27(1):25-37. [doi: 10.1016/S0305-0483(98)00028-0]
76. Andrews JC, Netemeyer RG, Burton S, Moberg DP, Christiansen A. Understanding adolescent intentions to smoke: an examination of relationships among social influence, prior trial behavior, and antitobacco campaign advertising. J Marketing 2004 Jul;68(3):110-123. [doi: 10.1509/jmkg.68.3.110.34767]
77. Cliff N. Analyzing multivariate data. San Diego: Harcourt Brace Jovanovich; 1987.

78. Winer B, Brown RB, Michels KM. Statistical principles in experimental design. New York: McGraw-Hill; 1971.
79. Davidson R, MacKinnon J. Estimation and inference in econometrics. New York: Oxford University Press; 1993.
80. Hsieh DA. A heteroscedasticity-consistent covariance matrix estimator for time series regressions. J Econometrics 1983 Aug;22(3):281-290. [doi: 10.1016/0304-4076(83)90104-5]
81. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [Medline: 3806354]
82. Holmbeck GN. Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: examples from the child-clinical and pediatric psychology literatures. J Consult Clin Psychol 1997 Aug;65(4):599-610. [Medline: 9256561]
83. Hoyle RH, Kenny DA. Sample size, reliability, and tests of statistical mediation. In: Hoyle RH, editor. Statistical strategies for small sample research. Thousand Oaks, CA: Sage; 1999:195-222.
84. Judd CM, Kenny DA. Process analysis: estimating mediation in treatment evaluations. Eval Rev 1981 Oct 01;5(5):602-619. [doi: 10.1177/0193841X8100500502]
85. Kline RB. Principles and Practice of Structural Equation Modeling. New York: The Guilford Press; 2015.
86. Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. J Couns Psychol 2004;51(1):115-134. [doi: 10.1037/0022-0167.51.1.115]
87. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
88. Schwender C. Evolutionary pychological components of a media theory. In: Hofer W, editor. Media and Emotions. Wiesbaden: Springer; 2013.
89. Meyer JC. Humor in member narratives: Uniting and dividing at work. Western J Comm 1997 Jun;61(2):188-208. [doi: 10.1080/10570319709374571]
90. Bartsch A, Appel M, Storch D. Predicting emotions and meta-emotions at the movies: The role of the Need for Affect in audiences? experience of horror and drama. Commun Res 2010 Feb 10;37(2):167-190. [doi: 10.1177/0093650209356441]
91. Biener L, Ji M, Gilpin EA, Albers AB. The impact of emotional tone, message, and broadcast parameters in youth anti-smoking advertisements. J Health Commun 2004;9(3):259-274. [doi: 10.1080/10810730490447084] [Medline: 15360037]
92. Sutfin EL, Szykman LR, Moore MC. Adolescents' responses to anti-tobacco advertising: exploring the role of adolescents' smoking status and advertisement theme. J Health Commun 2008;13(5):480-500. [doi: 10.1080/10810730802198961] [Medline: 18661389]
93. Lindley CA, Nacke L, Sennersten CC. Dissecting play-investigating the cognitive and emotional motivations and affects of computer gameplay. In: Proceedings of the International Conference on Computer Games (CGames 2008). 2008 Presented at: The International Conference on Computer Games (CGames 2008); 11/03/2008; Wolverhampton, UK p. 978.
94. Khalil GE. Change through experience: how experiential playemotional engagement drive health game success. In: Novak D, Tulu B, Bendryen H, editors. Public Health and Welfare: Concepts, Methodologies, Tools, and Applications. Hershey, PA: Medical Information Science Reference IGI Global; 2017:1421-1446.

Abbreviations

ANOVA: analysis of variance
ASPIRE: A Smoking Prevention Interactive Experience
CD: coefficient of determination
CFI: comparative fit index
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
E-ELM: extended-elaboration likelihood model
ELT: experiential learning theory
RMSEA: root mean square error of approximation
YMCA: Young Men's Christian Association

Edited by G Eysenbach; submitted 15.12.16; peer-reviewed by G Alexander, R Mermelstein; comments to author 11.01.17; revised version received 24.01.17; accepted 24.01.17; published 16.02.17.

Please cite as:
Khalil GE, Wang H, Calabro KS, Mitra N, Shegog R, Prokhorov AV
From the Experience of Interactivity and Entertainment to Lower Intention to Smoke: A Randomized Controlled Trial and Path Analysis of a Web-Based Smoking Prevention Program for Adolescents
J Med Internet Res 2017;19(2):e44
URL: http://www.jmir.org/2017/2/e44/
doi:10.2196/jmir.7174
PMID:28209560

Original Paper

The Benefit of Web- and Computer-Based Interventions for Stress: A Systematic Review and Meta-Analysis

Elena Heber[1,2,3], PhD; David Daniel Ebert[2,4], PhD; Dirk Lehr[1,2], PhD; Pim Cuijpers[2,5], PhD; Matthias Berking[2,4], PhD; Stephanie Nobis[2], PhD; Heleen Riper[2,5,6], PhD

[1]Department of Health Psychology and Applied Biological Psychology, Institute of Psychology, Leuphana University Lueneburg, Lueneburg, Germany
[2]Division of Online Health Training, Innovation Incubator, Leuphana University Lueneburg, Lueneburg, Germany
[3]Faculty of Social and Human Sciences, University of Southampton, Southampton, United Kingdom
[4]Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-University Erlangen Nuremberg, Erlangen, Germany
[5]Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[6]Telepsychiatric Centre, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Elena Heber, PhD
Department of Health Psychology and Applied Biological Psychology
Institute of Psychology
Leuphana University Lueneburg
Scharnhorststr 1
Lueneburg, 21335
Germany
Phone: 49 4131 677 ext 2720
Fax: 49 4131 677 1717
Email: heber@leuphana.de

Abstract

Background: Stress has been identified as one of the major public health issues in this century. New technologies offer opportunities to provide effective psychological interventions on a large scale.

Objective: The aim of this study is to investigate the efficacy of Web- and computer-based stress-management interventions in adults relative to a control group.

Methods: A meta-analysis was performed, including 26 comparisons (n=4226). Cohen *d* was calculated for the primary outcome level of stress to determine the difference between the intervention and control groups at posttest. Analyses of the effect on depression, anxiety, and stress in the following subgroups were also conducted: risk of bias, theoretical basis, guidance, and length of the intervention. Available follow-up data (1-3 months, 4-6 months) were assessed for the primary outcome stress.

Results: The overall mean effect size for stress at posttest was Cohen *d*=0.43 (95% CI 0.31-0.54). Significant, small effects were found for depression (Cohen *d*=0.34, 95% CI 0.21-0.48) and anxiety (Cohen *d*=0.32, 95% CI 0.17-0.47). Subgroup analyses revealed that guided interventions (Cohen *d*=0.64, 95% CI 0.50-0.79) were more effective than unguided interventions (Cohen *d*=0.33, 95% CI 0.20-0.46; *P*=.002). With regard to the length of the intervention, short interventions (≤4 weeks) showed a small effect size (Cohen *d*=0.33, 95% CI 0.22-0.44) and medium-long interventions (5-8 weeks) were moderately effective (Cohen *d*=0.59; 95% CI 0.45-0.74), whereas long interventions (≥9 weeks) produced a nonsignificant effect (Cohen *d*=0.21, 95% CI –0.05 to 0.47; *P*=.006). In terms of treatment type, interventions based on cognitive behavioral therapy (CBT) and third-wave CBT (TWC) showed small-to-moderate effect sizes (CBT: Cohen *d*=0.40, 95% CI 0.19-0.61; TWC: Cohen *d*=0.53, 95% CI 0.35-0.71), and alternative interventions produced a small effect size (Cohen *d*=0.24, 95% CI 0.12-0.36; *P*=.03). Early evidence on follow-up data indicates that Web- and computer-based stress-management interventions can sustain their effects in terms of stress reduction in a small-to-moderate range up to 6 months.

Conclusions: These results provide evidence that Web- and computer-based stress-management interventions can be effective and have the potential to reduce stress-related mental health problems on a large scale.

(J Med Internet Res 2017;19(2):e32) doi:10.2196/jmir.5774

KEYWORDS

stress; mental health; internet-based interventions; review; randomized controlled trial; meta-analysis

Introduction

Chronic stress can lead to serious psychological and physical implications, such as depression [1], sleep problems [2], neck and shoulder disorders [3], and an increased risk of coronary heart disease events [4] and related mortality [5]. Substantial economic costs of stress and stress-related psychological disorders arise as a result of absenteeism, presenteeism, productivity loss, and high staff turnover [6]. Given that stress represents a major threat to public health, effective and scalable solutions to accommodate the demand for stress-management interventions are needed.

The evidence base on traditional face-to-face stress-management interventions is comprehensive, showing small-to-moderate overall effects in reducing adverse outcomes for mental health [7]. In their meta-analysis on interventions for work-related stress, van der Klink et al [8] reported a combined effect size of Cohen *d*=0.34 across all studies. The mean effect size was Cohen *d*=0.33 for depression and Cohen *d*=0.54 for anxiety for interventions that focused on individuals [8]. Richardson and Rothstein's more recent meta-analysis [9] on occupational cognitive behavioral, relaxation, organizational, multimodal, or alternative stress-management interventions yielded somewhat larger effects, with an overall effect of Cohen *d*=0.53, an effect of Cohen *d*=0.68 for anxiety, and Cohen *d*=0.73 for stress. Evidence consistently shows cognitive behavioral interventions to be the most effective, with Cohen *d* values ranging from 0.68 [8] to 1.16 [9]. Some evidence also suggests that shorter interventions (1-4 weeks) are more effective than longer interventions; however, this effect may be confounded by treatment type [9].

A promising medium to facilitate the dissemination of effective mental health interventions is the Internet. In recent years, Web-based interventions for the prevention and treatment of a range of psychological conditions have proliferated [10-14]. Computer- and Web-based interventions are perceived to offer several advantages that may overcome some of the limitations of face-to-face approaches, including anonymity, 24/7 availability, reduced costs in terms of traveling to courses for both participants and instructors, high scalability, and a low access threshold. Enabling participants to be reached earlier than in classical face-to-face trainings, such interventions may have the potential to prevent the onset of more severe mental health problems [15-18]. Internet-based interventions have been found to be effective in clinical applications to address, for instance, depression [10], anxiety [19], and sleep disorders [20]. However, only a few randomized controlled trials (RCTs) have investigated Web-based stress-management interventions. Research on the eff0.19

Existing Web-based stress-management trainings differ in various aspects, such as the intervention content, length, and guidance, which may influence their efficacy. First, the theoretical basis of such interventions is diverse, including cognitive behavioral therapy (CBT) [21-26], third-wave cognitive behavioral interventions (TWC) [27-37], the use of olfactory care products [38], and physical exercise programs [39]. Second, the length and number of intervention sessions vary, ranging from short 2-week interventions [33] to interventions that allow access over several months [37]. Some interventions encourage participants to log on only as often as they like, with no specified sessions or requirements to complete the entire intervention [21], whereas others have fixed weekly appointments (eg, in an online virtual classroom) [37]. Third, the existing Web- and computer-based studies include both guided and unguided interventions and thus differ in the amount of human support given to participants during the intervention. In the guided format, individuals normally receive written feedback from a coach on the exercises that they complete within the training. In the study of Ruwaard et al [24], for example, clinical psychology students provided weekly feedback on exercises according to an instruction manual and reminded participants in cases of noncompletion. In contrast, Billings et al [21] used a less intensive, high-latitude format in which no feedback was provided.

In recent years, the number of studies on Web- and computer-based stress management has been rising. The overall effect of Web-based stress-management interventions and the influence of specific formats of treatment delivery remain unclear. Considering the demand for effective, scalable stress-management trainings and the enormous potential of Web- and computer-based interventions, there is a need to synthesize the results of existing studies. This meta-analysis aims to integrate the effects of Web-based stress-management interventions on the level of stress of adults. Additionally, effects on depression and anxiety will be assessed. The following research questions are addressed:

Are Web- and computer-based stress-management interventions effective in reducing stress, depression, and anxiety relative to a control group?

Are there differences in effect sizes concerning (a) the study quality, (b) the level of guidance, (c) the theoretical basis, and (d) the length of the intervention?

Methods

Eligibility Criteria

We considered RCTs from 1990 to May 2016 in which adult participants (older than 18 years) experienced stress and were participating in the trial to decrease their stress levels. The search was initially conducted in August 2013 and repeated in May 2016 to ensure it was as current as possible. Studies prior to 1990 were excluded; due to the rapidly changing technology in this field, these programs cannot be compared to the current interventions that are likely to be delivered to participants seeking help. The primary intervention target of included studies needed to be a reduction of stress within the target group. We excluded trials that targeted participants with medical conditions (eg, cancer, tinnitus, headache); caregivers of people with medical conditions (eg, caregivers for dementia); and participants with psychiatric disorders (eg, depression, anxiety or posttraumatic stress disorder), posttraumatic symptoms, postpartum emotional distress, or bereavement. The studies included needed to compare an intervention to any type of comparison group. There were no restrictions with regard to dosage or intensity, delivery, duration, frequency or timing of delivery, or the type of delivery channel (eg, email, Web-based, computerized). Trials investigating stress management as merely one part of the intervention (eg, alongside depression or anxiety) were included only if the primary goal of the intervention was to reduce stress. Furthermore, trials conducted in the context of the promotion of well-being rather than stress reduction were excluded. The studies had to report at least one instrument that claimed to measure stress levels. For the main analysis, we considered postintervention data. We also included follow-up data (when available) to assess longer-term effectiveness.

Information Sources and Search Strategy

The search strategy for this meta-analysis was created with four categories of search terms. We defined the search terms to meet the following criteria: (1) stress reduction, (2) evaluation of an intervention program, (3) application of a RCT as an experimental design, and (4) delivery in a Web- or computer-based context. A detailed description of the search terms can be found in the Multimedia Appendix 1.

The specified search strategy was applied in three major database search engines (PsycINFO, PubMed, and Cochrane). Additionally, manual searches in key journals (eg, *International Journal of Stress Management*, *Journal of Medical Internet Research*, *Journal of Occupational Health*, *Scandinavian Journal of Work, Environment and Health*) and in the reference lists of the included studies were conducted.

Study Selection

After removing duplicates of the articles identified, two researchers (EH and SN) independently screened all titles and abstracts based on the inclusion and exclusion criteria, and two researchers (EH and DL) assessed the full-text articles for eligibility. The researchers who assessed the relevance of the studies were not blind to the authors, institutions, journal of publication, or results.

Data Extraction

Data were extracted concerning the origin, the number of participants, age, gender, comparison group, outcomes, theoretical basis, length, guidance, the risk of bias, and follow-up data. In cases of insufficient description, the primary investigators of the respective studies were contacted to obtain missing information.

Risk of Bias

Two researchers (EH and DL) assessed the risk of bias of the included studies in accordance with the Cochrane Guidelines [40]. Thereby, sequence generation, allocation concealment, blinding, incomplete outcome data, and selective outcome reporting were judged.

Power Calculation

Because we expected only a limited number of studies with relatively small sample sizes, we conducted a power calculation to examine how many studies with how many participants would need to be included to assure sufficient statistical power to identify relevant effects. This power calculation was conducted according to the procedures described by Borenstein and colleagues [41]. We hoped to find a sufficient number of studies to be able to identify a small-to-moderate effect size of 0.35. These calculations indicated that we would need to include at least five studies with a mean sample size of 100 (50 participants per condition) to be able to detect an effect size of Cohen *d*=0.35 (conservatively assuming a medium level of between-study variance, T^2, a statistical power of .80, and a significance level, alpha, of .05). Alternatively, we would need three studies with 150 participants each to detect an effect size of Cohen *d*=0.35 or nine studies with 50 participants.

Analyses

Analyses were conducted using the statistical software Comprehensive Meta-Analysis (version 2.2.057). The effect size of subjective level of stress was calculated as the primary outcome. Furthermore, we assessed the levels of depression and anxiety when available. If more than one measure per outcome was used and if no primary outcome was indicated, then the mean of the effect size was calculated to ensure that each study yielded only one effect size. A random-effects model was chosen because of the expectations of considerable heterogeneity between studies. We further conducted a series of subgroup analyses according to the mixed-effects model. In this model, studies within subgroups are pooled with the random-effects model, whereas tests for significant differences between subgroups are conducted with the fixed-effects model. Because of the small number of studies providing follow-up data, the subgroup analyses were performed only for posttreatment and the results from the follow-ups for the primary outcome stress were clustered into two categories (1-3 months and 4-6 months).

In three studies, two treatments were compared with a single control group [36,38,42]. In these cases, we treated each comparison as a separate study, and we avoided the double counting of controls by dividing the control group in half.

The effect size in the form of Cohen *d* was used to represent the standard mean difference between the means of the intervention group and the control group at posttest. According to Cohen [43], *d*=0.2 can be considered a small effect, *d*=0.5 is a medium effect, and *d*=0.8 is a large effect. Because of the difficulty of interpreting Cohen *d* from a clinical perspective, we also transformed these values into numbers needed to treat (NNT) according to the formula of Kraemer and Kupfer [44]. The NNT indicates the number of participants who need to be treated to generate one additional clinically significant change [45].

Furthermore, we conducted the following subgroup analyses: the theoretical basis of the intervention (CBT, identified by including cognitive restructuring/challenging dysfunctional thoughts; TWC, identified by inclusion of more recent CBT-based techniques such as mindfulness, meditation, or acceptance of emotions; and alternative interventions [ALT]), guidance (guided with regular written feedback; unguided with no support or only technical support), length of the intervention (short: 1-4 weeks; medium: 5-8 weeks; long: ≥9 weeks), and the risk of bias (low risk=4; some risk<4).

Moreover, a test of homogeneity of the observed effect sizes was calculated using the I^2 statistic as an indicator of heterogeneity in percentages. Thereby, a value of 0% indicates no heterogeneity, 25% is considered low, 50% is considered moderate, and 75% is considered a high level of heterogeneity [46]. We calculated 95% confidence intervals around I^2 [47] using the noncentral approach based on chi-square within the heterogi module for Stata [48]. Although we calculated the Q-statistics, we only report whether the result was significant.

Publication bias was investigated by conducting a visual inspection of the funnel plot for the primary outcome measure. An asymmetric as opposed to a symmetric inverted funnel shape indicates potential publication bias that could compromise the conclusions drawn from the meta-analysis. Egger's test [49] was used to test for the significance of the likely presence of publication bias. Additionally, we performed Duval and Tweedie's trim-and-fill analysis [50] to verify an unbiased estimate of the pooled effect size. This analysis calculates an estimation of the number of missing studies and the potential effect of these studies on the outcome.

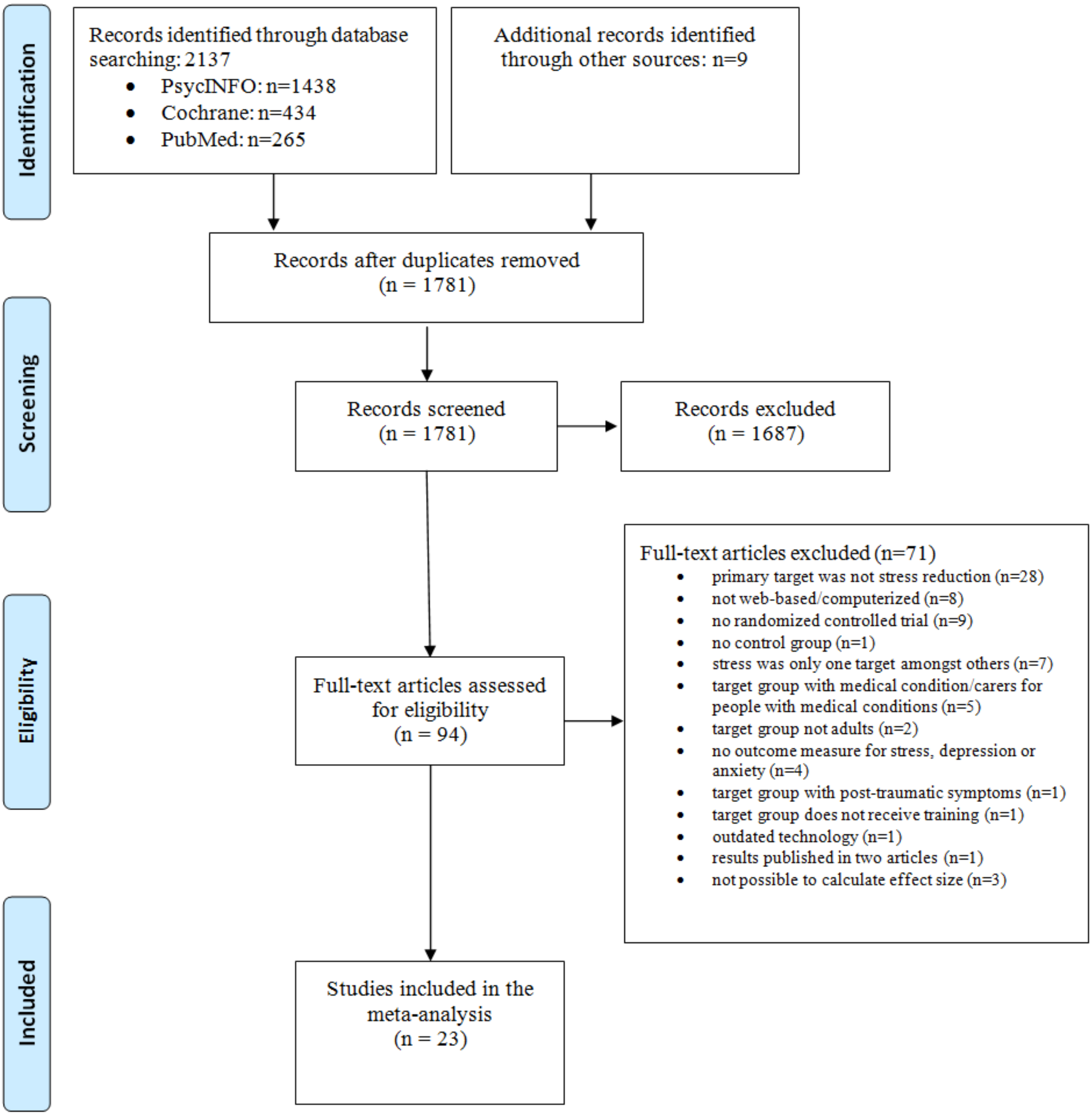
Results

Study Selection

The systematic literature research resulted in 2137 abstracts. An additional nine potentially relevant articles were identified through other searches. After removing the duplicates, we screened the titles and abstracts of 1781 articles and excluded 1687 articles because of their apparent irrelevance. With regard to eligibility, 94 full-text articles were retrieved and assessed by two independent raters (EH and DL); Cohen kappa for agreement between the independent raters was very good (Cohen κ=.83). Any discrepancies were resolved by discussion. We included 27 studies according to the inclusion and exclusion criteria. However, the results of one study [51] were accounted for in one of the other included articles [24] and it was not possible to calculate effect sizes for three studies due to insufficient data [52-54]. In three studies [42,55,56], a small proportion of the participants were younger than 18 years, which was the prespecified criterion of being classified as adults. Because the studies fit all the other inclusion criteria and our sensitivity analyses indicated no difference in the overall results, we decided to include those studies.

Thus, we included 23 studies in the analysis. These 23 studies included 26 comparisons from baseline to posttest. Follow-up data were available for four studies (six comparisons) at 1 to 3 months and for six studies at 4 to 6 months. The PRISMA flowchart of the study selection is presented in Figure 1.

Figure 1. PRISMA flowchart.



Study Characteristics

Selected characteristics of the 23 included studies [21-39,42,55-57] are presented in Table 1 (trial characteristics) and Table 2 (intervention characteristics). A more detailed description of the interventions is available in the Multimedia Appendix 2.

The total number of participants was 4226 (intervention groups: n=2312, control groups: n=1914). The included interventions varied according to the intervention content (see Table 2 for details). Most studies evaluated interventions based on TWC (13 comparisons), followed by ALT interventions (7 comparisons), and CBT interventions (6 comparisons). Seventeen studies used a waitlist control comparison, three studies a no-treatment control group, and three studies used an attention control group. The included studies were predominantly conducted in the United States (11 studies), followed by Germany (3 studies), Sweden (2 studies), Austria/Switzerland (2 studies), Japan (2 studies), Norway (1 study), the United Kingdom (1 study), and the Netherlands (1 study).

Table 1. Selected trial characteristics of the included studies (N=23).

Study	Origin[a]	N[b]	Age (years), mean/mean range	Women (%)	Conditions[c]	Outcomes[d]	Risk of bias[e]	Follow-up[f]
Aikens et al (2014) [27]	US	89	N/A	N/A	TWC vs WC	Stress (PSS-14)	+? ± ??	6 m ext
Allexandre et al (2016) [28]	US	55	40.0	83.2	TWC vs WC	Stress (PSS-10)	+? ± +?	4 m; 1 y ext
Billings et al (2008) [21]	US	236	30-39	70.6	CBT vs WC	Stress (SODS); depression (CES-D); anxiety (BAI)	+? ± – ?	N/A
Cavanagh et al (2013) [29]	GB	104	24.7	88.5	TWC vs WC	Stress (PSS-10); depression (PHQ-4); anxiety (PHQ-4)	+? ± +?	N/A
Drozd et al (2013) [30]	NO	259	32.6	76.0	TWC vs WC	Stress (DASS-S)	+ + ± +?	2 m; 6 m
Ebert et al (2016) [31]	DE	263	42.0	71.5	TWC vs WC	Stress (PSS-10); depression (CES-D); anxiety (HADS-A)	+? ± + +	6 m
Ebert et al (2016) [32]	DE	263	42.9	85.9	TWC vs WC	Stress (PSS-10); depression (CES-D) anxiety (HADS-A)	+? ± + +	6 m
Frazier et al (2015) [56]	US	194	18-21	75.0	ALT vs AC	Stress (PSS-10, DASS-S); depression (DASS-D) anxiety (DASS-A)	+? ± +?	5 w
Glück & Maercker (2011) [33]	AT, CH	47	35.2	73.5	TWC vs WC	Stress (PSQ)	+ – ± +?	3 m ext
Hänggi (2006) [22]	CH	125	25-45	85.0	CBT vs WC	Stress (LOS)	+ + ± – ?	1 m ext; 6 m ext
Heber et al (2016) [34]	DE	264	43.3	73.1	TWC vs WC	Stress (PSS-10); depression (CES-D); anxiety (HADS-A)	+? ± + +	6 m; 1 y ext
Hinman et al (1997) [39]	US	50	37.7	100.0	ALT vs NT	Stress (PNQ (combined)	+? ± ? ?	N/A
Hintz et al (2015) [42]	US	204	18-21	70.0	ALT vs AC; ALT vs AC (II)	Stress (PSS-10, DASS-S); depression (DASS-D); anxiety (DASS-A)	+? ± – ?	5 w
Ly et al (2014) [35]	SE	73	41.5	42.5	TWC vs WC	Stress (PSS-14)	+ + ± + ?	N/A
Morledge/ Allexandre et al (2013) [36]	US	279	40-59	88.9	TWC vs TWC+OMB vs WC	Stress (PSS-10)	+ + ± – ?	1 m
Nguyen-Feng et al (2015) [55]	US	500	18-21	62.0	ALT vs WC	Stress (PSS-10, DASS-S); depression (DASS-D); anxiety (DASS-A)	+ ? ± + ?	N/A
Rose et al (2013) [23]	US	59	27.3	50.0	CBT vs AC	Stress (PSS-10)	+ ? ± – ?	N/A
Ruwaard et al (2007) [24]	NL	239	44.0	60.0	CBT vs WC	Stress (DASS-S); depression (DASS-D); anxiety (DASS-A)	+ ? ± + ?	3 y ext

Study	Origin[a]	N[b]	Age (years), mean/mean range	Women (%)	Conditions[c]	Outcomes[d]	Risk of bias[e]	Follow-up[f]
Umanodan et al (2014) [25]	JP	263	38.9	7.2	CBT vs WC	Stress (BJSQ)	+ ? ± + ?	19 w
Wiegand et al (2010) [38]	US	562	35.8	100.0	ALT vs TWC vs NT	Stress (PSS-10); anxiety (STAI)	+ ? ± – ?	N/A
Wolever et al (2012) [37]	US	105	42.9	76.6	TWC vs NT	Stress (PSS-10); depression (CES-D)	+ + ± + ?	N/A
Yamagishi et al (2008) [57]	JP	36	33.0	N/A	ALT vs WC	Stress (JSBQ combined); depression (BJSQ-D); anxiety (BJSQ-A)	+ ? ± – ?	1 m ext
Zetterqvist et al (2003) [26]	SE	63	39.2	61.9	CBT vs WC	Stress (PSS-14); depression (HADS-D); anxiety (HADS-A)	+ ? ± – ?	N/A

[a] AT: Austria; CH: Switzerland; DE: Germany; GB: United Kingdom; JP: Japan; NL: the Netherlands; NO: Norway; SE: Sweden; US: United States of America.

[b] Indicates only the number of participants included in this analysis.

[c] AC: attention control group; ALT: alternative; CBT: cognitive behavioral therapy; NT: no treatment; OMB: online message board; TWC: third-wave cognitive behavioral therapy; WC: waitlist control.

[d] BAI: Beck Anxiety Inventory; BJSQ-A: Brief Job Stress Questionnaire-Anxiety Subscale; BJSQ-D: Brief Job Stress Questionnaire-Depression Subscale; CES-D: Center for Epidemiologic Studies Depression Scale; DASS-A: Depression Anxiety Stress Scales-Anxiety Subscale; DASS-D: Depression Anxiety Stress Scales-Depression Subscale; DASS-S: Depression Anxiety Stress Scales-Stress Subscale; HADS-A: Hospital Anxiety and Depression Scales-Anxiety Subscale; HADS-D: Hospital Anxiety and Depression Scales-Depression Subscale; JSBQ: Job Stress Brief Questionnaire; LOS: Level of Stress (self-created instrument); PHQ-4: Patient Health Questionnaire for Depression and Anxiety; PNQ: Personal Strain Questionnaire; PSQ: Perceived Stress Questionnaire; PSS-10: Cohen’s Perceived Stress Questionnaire (10-item version); PSS-14: Cohen’s Perceived Stress Questionnaire (14-item version); SODS: Symptoms of Distress Scale; STAI: State Trait Anxiety Inventory.

[e] Risk of bias was judged according to the following criteria: (1) adequate sequence generation, (2) allocation concealment, (3) blinding (± indicates that only self-reported data were used), (4) adequate consideration of incomplete data, and (5) prevention of selective outcome. +: no bias; –: bias; ?: information was insufficient to make judgments.

[f] Ext: extended follow-up; m: months; w: weeks; y: years.

Table 2. Selected intervention characteristics of included studies.

Study	Label and content	Type[a]	Guidance[b]	Delivery	Length (weeks)
Aikens et al (2014) [27]	Mindfulness goes to work: Mindfulness program combined live, instructor-led, weekly hour-long virtual meetings (webinar) with online applied training.	TWC	G	Web	7
Allexandre et al (2016) [28]	Stress Free Now / Online Mindfulness Program for Stress Management: Interactive, educational program based on mindfulness meditation. Includes exercises, email reminders, and downloads.	TWC	UG (R)	Web	8
Billings et al (2008) [21]	Stress and Mood Management Intervention: Array of CBT techniques. Entire program is audio-narrated with the use of videos and graphics.	CBT	UG	Web	12
Cavanagh et al (2013) [29]	Learning Mindfulness Online: Daily, 10-min guided mindfulness meditation audio tracks delivered via a virtual learning facility (Moodle). Four reminder emails.	TWC	UG (R)	Web	2
Drozd et al (2013) [30]	Less Stress intervention: Eclectic approach that included mindfulness and metacognitive exercises with 13 short modules. Hyperlinks sent via email to the participants.	TWC	UG	Web	4
Ebert et al (2016) [31]	GET.ON Stress Self-Guided: Theory-based intervention focusing on problem solving and emotion regulation. Tailored to employees; optional text message coaching.	TWC	UG	Web	7
Ebert et al (2016) [32]	GET.ON Stress Adherence-Focused Guided: Theory-based intervention focusing on problem solving and emotion regulation. Tailored to employees; optional text message coaching; written feedback on request; reminders.	TWC	AFG	Web	7
Frazier et al (2015) [56]	Present Control Intervention: Theory-based intervention focused on perceived control; 4 modules over 2-week period. Included stress logs and reminder emails.	ALT	UG (R)	Web	2
Glück & Maercker (2011) [33]	Brief Web-based mindfulness training: 2 modules. 20-minute units per day, audio files, a flash animated exercise, and written text.	TWC	UG (R)	Web	2
Hänggi (2006) [22]	Online parental training on coping with family stress: 4 modules (eg, cognitive restructuring, time management, muscle and breathing relaxation, problem solving).	CBT	UG	Web	4
Heber et al (2016) [34]	GET.ON Stress Guided: Theory-based intervention focusing on problem solving and emotion regulation. Tailored to employees; optional text message coaching; written feedback; reminders.	TWC	G	Web	7
Hinman et al (1997) [39]	Exercise Break: 2 × 15 minutes per day. Stretching, circulatory and relaxation exercises. Accessed via local computer network at the workplace.	ALT	UG	PC	8
Hintz et al (2015) [42]	Present Control Intervention: Theory-based intervention focused on perceived control; 4 modules over 2-week period. Included stress logs and reminder emails. Group I: with personalized feedback, Group II: unguided.	ALT	UG (R) & G	Web	2
Ly et al (2014) [35]	Acceptance- and commitment-based mobile phone app: step-by-step behavior program including 6 basic principles to handle stress. 15 min daily. Short writing reflection. One-way therapist-client support through text messages every other day.	TWC	G	Web (smartphone)	6
Morledge/Allexandre et al (2013) [36]	Online Mindfulness Program for Stress Management: Eight mindfulness modules consisting of introduction, meditations, articles, and tips and exercises. Group II: program plus online message board.	TWC	UG (R)	Web	8
Nguyen-Feng et al (2015) [55]	Present Control Intervention: Theory-based intervention focused on perceived control; 3 modules: psycho-educational video of a professor, animated video (Prezi), and a written exercise. Includes stress logs and reminders.	ALT	UG (R)	Web	5
Rose et al (2013) [23]	Self-guided, multimedia stress management and resilience training program, SMART-OP: consists of at least one exercise in each of 3 domains: feelings, thoughts, and actions. Includes game-like activities.	CBT	UG (R)	PC	6
Ruwaard et al (2007) [24]	Emailed Standardized CBT of Work-Related Stress: 7 modules (eg, relaxation, challenging dysfunctional thoughts, time management). 10 feedbacks/5 hours of therapist time.	CBT	G	Web	7
Umanodan et al (2014) [25]	SMT program in employees: Self-paced program. (1) behavioral techniques, (2) communication techniques, and (3) cognitive techniques; skill acquisition and practice phase; weekly emails.	CBT	UG (R)	PC	7
Wiegand et al (2010) [38]	Comprehensive program for reducing stress: Group I: Daily use of olfactory care products plus an Internet-based program focusing on stress reduction, prevention and behavioral modification. Periodic feedback reports are provided. Group II: Internet-based program only.	ALT, TWC	UG	Web	12

Study	Label and content	Type[a]	Guidance[b]	Delivery	Length (weeks)
Wolever et al (2012) [37]	Mindfulness at Work Intervention: Virtual classroom with real-time bidirectional communication with an experienced mindfulness trainer (12 modules, 14 hours). Includes brief exercises designed to be used at work.	TWC	G	Web	12
Yamagishi et al (2008) [57]	Web-based career identity training for stress management: 4 modules. (1) definition of career identity, (2): cognition of own career identity, (3): characteristics of nurses' career identity, (4): career goal management and planning.	ALT	UG	Web	3
Zetterqvist et al (2003) [26]	Internet-based self-help stress-management program: Each module consists of 3 sections: relaxation, additional exercises (eg, problem solving), and information (eg stress at work). Exercises were sent in and participants received feedback as a prompt to continue; includes reminders.	CBT	G	Web	6

[a] ALT: alternative; CBT: cognitive behavioral therapy; TWC: third-wave cognitive behavioral therapy.

[b] AFG: Adherence-focused guidance; G: guided; UG: unguided; UG (R): unguided with reminders via mail or telephone.

The interventions of the included studies were mainly Web-based interventions (n=20). For Web-based interventions, an active Internet connection is necessary (eg, to access a website, use a mobile phone app, or visit a virtual classroom). A total of three computer-based interventions [23,25,39] were found in which interventions were installed, for example, on a desktop computer in a separate room at work.

The Perceived Stress Scale (PSS-10, PSS-14) was predominantly used to assess the level of stress (15 of 23 studies). Follow-up assessments have been reported for nine studies (12 comparisons); whereas four studies (6 comparisons) reported data for up to 3 months and six studies (6 comparisons) for 4 to 6 months (see Table 1). Extended follow-ups (no comparison with the respective control groups) were conducted at 1 month (2 studies), 3 months (1 study), 6 months (2 studies), one year (2 studies), and 3 years (1 study). For nine studies, only posttest assessments were available. The participants received guidance in seven studies [24,26,27,34,35,37,42], one study [32] assessed a less intensive guidance format (adherence-focused guidance: combination of reminders and written feedback only on request of the participants), and 16 studies (18 comparisons) investigated unguided interventions. Nine unguided studies reported the use of automated or telephone and mail reminders for completion of the intervention [23,25,28,29,33,36,42,55,56] (Table 2).

Risk of Bias Within Studies

A risk of bias was present in all studies. Only six studies fulfilled four of the five criteria used. Nine studies met three criteria, and eight studies fulfilled two criteria. In most cases, the concealment of allocation was insufficiently described, and only 13 of 23 studies reported adequate handling of missing data. In particular, the risk for selective outcome reporting was unclear because the study registration prior to the trials could only be retrieved from three studies [31,32,34]. Although another four studies registered their trial [28,30,33,36], this step occurred retrospectively. The mean interrater reliability between independent raters was κ=.84 and ranged from .60 (selective outcome reporting) to 1.0 (blinding). Disagreements were handled by discussion.

Publication Bias

Neither the inspection of the funnel plot nor the Egger's test [49] indicated a possible publication bias. Moreover, the Duval and Tweedie trim-and-fill analysis [50] indicated no missing studies.

Effects on Levels of Stress, Depression, and Anxiety

Table 3 presents the effect sizes, confidence intervals, level of significance and heterogeneity for the overall effects on stress, depression, and anxiety as well as for the subgroups. The overall analysis of effect sizes yielded a significant effect size of Cohen *d*=0.43 for the primary outcome stress across all studies (95% CI 0.31-0.54; n=26). Heterogeneity was moderate (I^2=68.01, 95% CI 52.08-78.72). Significant small effect sizes were observed for the secondary outcomes depression (Cohen *d*=0.34, 95% CI 0.21-0.48; n=13) and anxiety (Cohen *d*=0.32, 95% CI 0.17-0.47; n=14). Figure 2 displays a forest plot of the effect sizes and the confidence intervals of the included studies.

Table 3. Effects of computer- and Web-based stress-management interventions for healthy adults compared to control groups.

Study	Comparisons, n	Cohen *d* (95% CI)	*P*	*Z*	I^2 (95% CI)	*P* [a]	NNT	*P* [b]
Overall effect								
Primary outcome								
Stress at posttest	26	0.43 (0.31, 0.54)	<.001	7.12	68.01 (52.08, 78.72)	<.001	4.20	
1-3 m follow-up	6	0.33 (0.19, 0.46)	<.001	4.60	0.00 (0.00, 74.62)	.55	5.43	
4-6 m follow-up	6	0.56 (0.25, 0.87)	<.001	3.55	85.93 (71.44, 93.07)	<.001	3.25	
Further outcomes (posttest)								
Depressive symptoms	13	0.34 (0.21, 0.48)	<.001	4.93	58.25 (22.81, 77.41)	.004	5.26	
Anxiety symptoms	14	0.32 (0.17, 0.47)	<.001	4.16	71.13 (50.34, 83.22)	<.001	5.56	
Risk of bias score [c]								
Low risk	6	0.74 (0.59, 0.89)	<.001	9.82	35.44 (0.00, 74.20)	.17	2.50	<.001
Some risk	20	0.30 (0.21, 0.40)	<.001	6.26	31.43 (0.00, 60.17)	.09	5.95	
Theoretical basis [c]								
CBT	6	0.40 (0.19, 0.61)	<.001	3.75	52.68 (0.00, 81.12)	.06	4.50	.03
TWC	13	0.53 (0.35, 0.71)	<.001	5.67	74.50 (55.98, 85.23)	<.001	3.42	
ALT	7	0.24 (0.12, 0.36)	<.001	4.03	0.00 (0.00,70.81)	.85	7.46	
Guidance [c,d]								
Yes	7	0.64 (0.50, 0.79)	<.001	8.53	11.81 (0.00, 74.24)	.34	2.86	.002
No	18	0.33 (0.20,0.46)	<.001	5.02	62.72 (38.03, 77.57)	<.001	5.43	
Length [c]								
Short	9	0.33 (0.22, 0.44)	<.001	5.94	0.00 (0.00, 64.80)	.56	5.43	.006
Medium	13	0.59 (0.45, 0.74)	<.001	7.89	54.92 (15.83, 75.86)	.008	3.09	
Long	4	0.21 (-0.05, 0.47)	.11	1.61	71.00 (17.25, 89.85)	.02	8.47	

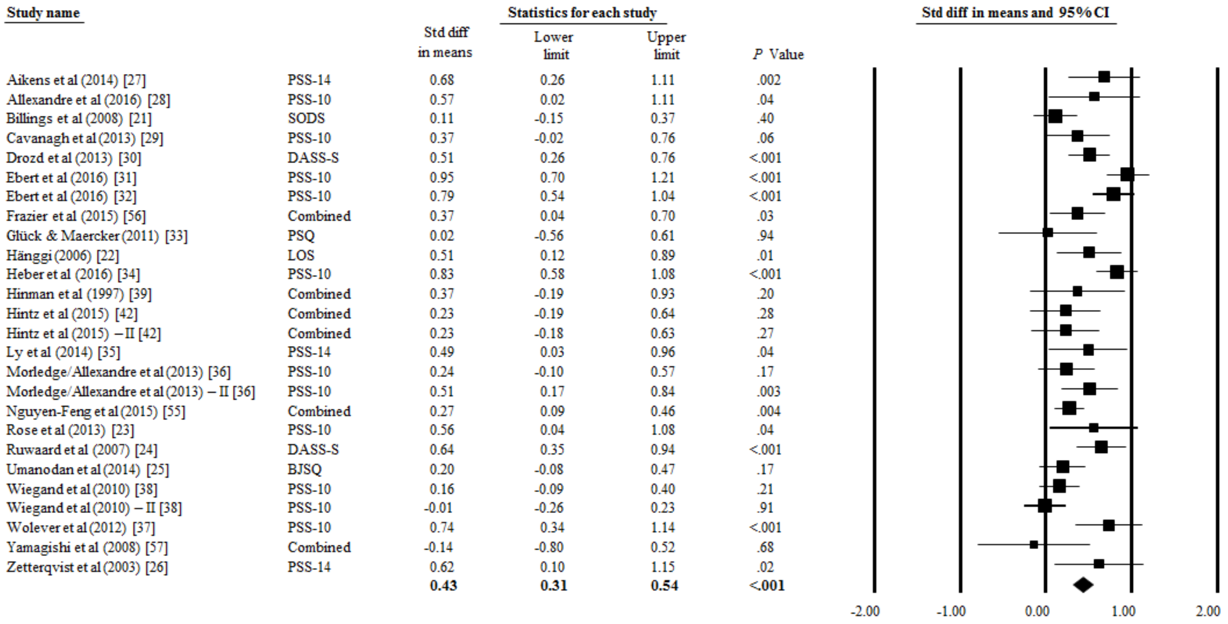
[a] The *P* values indicate whether the Q-statistic is significant (the I^2 statistics do not include a test of significance).

[b] This *P* value indicates whether the differences between subgroups were significant.

[c] Subgroup calculations for the primary outcome stress.

[d] One study [32] was excluded because it could not be classified.

Figure 2. All effect sizes included in the meta-analysis from the studies comparing computer- and Web-based stress-management interventions to a control group. BJSQ: Brief Job Stress Questionnaire; DASS-S: Depression Anxiety Stress Scales-Stress Subscale; LOS: Level of Stress (self-created instrument); PSQ: Perceived Stress Questionnaire; PSS-10: Cohen's Perceived Stress Questionnaire (10 item version); PSS-14: Cohen's Perceived Stress Questionnaire (14 item version); SODS: Symptoms of Distress Scale.



Study name		Statistics for each study			
		Std diff in means	Lower limit	Upper limit	*P* Value
Aikens et al (2014) [27]	PSS-14	0.68	0.26	1.11	.002
Allexandre et al (2016) [28]	PSS-10	0.57	0.02	1.11	.04
Billings et al (2008) [21]	SODS	0.11	-0.15	0.37	.40
Cavanagh et al (2013) [29]	PSS-10	0.37	-0.02	0.76	.06
Drozd et al (2013) [30]	DASS-S	0.51	0.26	0.76	<.001
Ebert et al (2016) [31]	PSS-10	0.95	0.70	1.21	<.001
Ebert et al (2016) [32]	PSS-10	0.79	0.54	1.04	<.001
Frazier et al (2015) [56]	Combined	0.37	0.04	0.70	.03
Glück & Maercker (2011) [33]	PSQ	0.02	-0.56	0.61	.94
Hänggi (2006) [22]	LOS	0.51	0.12	0.89	.01
Heber et al (2016) [34]	PSS-10	0.83	0.58	1.08	<.001
Hinman et al (1997) [39]	Combined	0.37	-0.19	0.93	.20
Hintz et al (2015) [42]	Combined	0.23	-0.19	0.64	.28
Hintz et al (2015) – II [42]	Combined	0.23	-0.18	0.63	.27
Ly et al (2014) [35]	PSS-14	0.49	0.03	0.96	.04
Morledge/Allexandre et al (2013) [36]	PSS-10	0.24	-0.10	0.57	.17
Morledge/Allexandre et al (2013) – II [36]	PSS-10	0.51	0.17	0.84	.003
Nguyen-Feng et al (2015) [55]	Combined	0.27	0.09	0.46	.004
Rose et al (2013) [23]	PSS-10	0.56	0.04	1.08	.04
Ruwaard et al (2007) [24]	DASS-S	0.64	0.35	0.94	<.001
Umanodan et al (2014) [25]	BJSQ	0.20	-0.08	0.47	.17
Wiegand et al (2010) [38]	PSS-10	0.16	-0.09	0.40	.21
Wiegand et al (2010) – II [38]	PSS-10	-0.01	-0.26	0.23	.91
Wolever et al (2012) [37]	PSS-10	0.74	0.34	1.14	<.001
Yamagishi et al (2008) [57]	Combined	-0.14	-0.80	0.52	.68
Zetterqvist et al (2003) [26]	PSS-14	0.62	0.10	1.15	.02
		0.43	**0.31**	**0.54**	**<.001**

Subgroup Analyses

Risk of Bias

Accounting for the risk of bias level was associated with a considerable reduction of heterogeneity: studies with a lower risk of bias showed low heterogeneity (I^2=35.44, 95% CI 0.00-74.20). Subgroup analyses revealed that studies at low risk produced significantly larger effect sizes (Cohen *d*=0.74, 95% CI 0.59-0.89; n=6) than did studies with some risk (Cohen *d*=0.30, 95% CI 0.21-0.40; n=20; *P*<.001).

Theoretical Basis

The subgroup analysis of the theoretical basis of the interventions was significant (*P*=.03) and showed that TWC interventions produced a highly significant medium effect size of Cohen *d*=0.53 (95% CI 0.35-0.71; n=13). CBT interventions led to reductions in stress levels with a highly significant effect size of Cohen *d*=0.40 (95% CI 0.19-0.61; n=6). In contrast, ALT interventions produced a small effect size (Cohen *d*=0.24; 95% CI 0.12-0.36; n=7).

Guidance

With regard to the subgroup guidance, the results show that guided interventions (Cohen *d*=0.64, 95% CI 0.50-0.79; n=7) were significantly more effective than unguided interventions (Cohen *d*=0.33, 95% CI 0.20-0.46; n=18; *P*=.002).

Length

A significant difference was also found for the length of the interventions (*P*=.006). We found significant small-to-medium effect sizes for short (Cohen *d*=0.33, 95% CI 0.22-0.44; n=9) and medium interventions (Cohen *d*=0.59, 95% CI 0.45-0.74; n=13), whereas long interventions led to a nonsignificant effect size of Cohen *d*=0.21 (95% CI –0.05 to 0.47; n=4; *P*=.11).

Follow-up

Results of the 1 to 3 month follow-ups for the primary outcome stress showed a small effect size of Cohen *d*=0.33 (95% CI 0.19-0.46; n=6) and, for the 4 to 6 month follow-ups, a medium effect size of Cohen *d*=0.56 (95% CI 0.25-0.87; n=6).

Discussion

The aim of this paper was to conduct a meta-analysis of Web- and computer-based interventions for stress management in adults. Our analyses yielded four main findings. First, Web- and computer-based stress-management interventions can be effective in reducing stress, depression, and anxiety and maintain those effects for up to 6 months. Second, interventions using TWC and CBT interventions as a theoretical basis proved moderately effective in reducing stress. Third, short and medium interventions (up to 8 weeks) are more effective than long interventions (9 weeks and longer). Fourth, guided interventions yielded a greater effect size for reduction of stress than unguided interventions.

For the primary outcome stress, an effect size of Cohen *d*=0.43 was found across the 26 comparisons. Small effect sizes were found for depression (Cohen *d*=0.34) and anxiety (Cohen *d*=0.32). A recent synthesis of five meta-analyses of traditional stress-management interventions found a between-group overall mean effect size of Cohen *d*=0.45 (95% CI 0.41-0.48) [58], which is comparable to the effect of stress, but somewhat higher for the effects of depression and anxiety found in this meta-analysis. To date, no trials have been conducted that aim to assess the equivalence of face-to-face and Web-based stress-management interventions in a methodologically robust design. One trial comparing an online versus face-to-face version of stress management in a single trial indicates that there is no difference in reductions of stress or depression levels [37].

Nevertheless, higher effect sizes were also found in Richardson and Rothstein's work [9] concerning individual outcome measures in traditional stress-management interventions, particularly for stress (Cohen *d*=0.73) and anxiety (Cohen *d*=0.68). Therefore, it may be possible that traditional interventions yield slightly higher effect sizes. One possible explanation for this is that face-to-face interventions are superior in reducing these outcomes. An alternative explanation is that Web-based interventions may reach affected individuals at an earlier stage, with lower baseline levels and thus less room for improvement. In fact, most studies included in the present work did not use a cut-off on a relevant stress scale, and most of those that did (eg, [23,37,38]) used a relatively low cut-off threshold. It may be the case that participants who are more severely impaired might generally prefer face-to-face over Web-based interventions and that this effect is reflected in greater improvements stemming from their higher baseline stress levels. This is in line with the fact that the highest effect sizes in this meta-analysis were produced from trials targeting highly stressed individuals [31,32,34]. Evidence on a Web-based depression intervention also showed that a higher severity of baseline scores significantly predicted better treatment outcomes [59]. Although the effect sizes found here are somewhat smaller than those found in traditional face-to-face interventions, Web- and computer-based interventions can have greater reach. At the population level, even small-to-moderate effects can have a substantial influence. More research is needed to clarify the effects of face-to-face and Web-based stress-management interventions in direct comparisons.

Based on the assumption that the treatment effect varied as a function of other factors, we conducted a number of subgroup analyses. First, we investigated the effect of guidance. The results showed that guided interventions (Cohen *d*=0.64) are significantly more effective than unguided interventions (Cohen *d*=0.33), with effect sizes for guided interventions comparable to traditional face-to-face interventions. The finding of guided interventions being superior to unguided interventions is consistent with results on Web-based interventions for other conditions, such as depression and anxiety [60,61]. Providing support to clients in terms of weekly feedback may enhance adherence to the intervention and thus improve treatment efficacy [62]. Such an assumption is in line with a study on pooled data from three RCTs showing that guidance was associated with greater adherence rates in Internet-based stress-management compared to unguided interventions [63]. The finding that guided interventions are more effective than unguided interventions is consistent with a systematic review that found guided Web-based mental health interventions to be significantly superior to unguided interventions [61]. The results in this analysis indicate an advantage for guided interventions, although it is unclear how much guidance and in what manner it produces the largest effect sizes. One trial in this meta-analysis used an alternative, more economic format of guidance (adherence-focused guidance; ie, feedback only on request plus weekly reminders) and produced large effect sizes at posttest and 6-month follow-up [32], and the adherence rates were comparable to a more intensive guidance format using the same intervention [63]. More research on the relative level and type of human involvement needed in these interventions would be useful [64].

Consistent with existing evidence on face-to-face interventions [8,9], the CBT interventions included in this study were efficacious. The effect sizes of CBT (Cohen *d*=0.40) and TWC interventions (Cohen *d*=0.53) were smaller than the average effect size found in two meta-analyses for traditional CBT interventions (Cohen *d*=0.68 [8]; Cohen *d*=1.16 [9]), although, on average, the effect sizes fall within the confidence interval of the latter work [9]. Direct comparisons examining the relative efficacy of the two training formats would be needed to draw firm conclusions. Due to their proliferation in the last few years, TWC interventions were introduced as a new category in this meta-analysis as they extend the traditional CBT techniques with newer “third-wave” components, such as acceptance of emotions or mindfulness. These interventions have been found to be effective in alleviating symptoms that are associated with a wide range of physical, psychosomatic, and psychiatric disorders [65,66], including stress [67]. Comparable to our findings, early evidence on face-to-face interventions suggests that TWC and traditional CBT approaches are equally effective and acceptable in the treatment of acute depression [68]; nevertheless, more high-quality studies are needed to support this assumption. As opposed to (third-wave) CBT interventions, alternative approaches (eg, career identity training, combination with olfactory care products) only produced a small effect size.

A significant between-group effect was also found for the length of the intervention. In contrast to short-to-medium interventions, long interventions (9 weeks and longer) were not found to be effective. One possible explanation is that it may be more difficult for participants to remain engaged in longer interventions compared with shorter interventions. These results correspond to findings from the area of depression, in which shorter interventions have been found to be more effective than longer interventions [61]. Research on the relevance of treatment intensity suggests that the number of therapy sessions is not related to the outcome and keeping the number of sessions equal, but providing the sessions over a shorter period of time, may be associated with better treatment outcomes [69]. Nevertheless, because there were only four comparisons available for long interventions, this conclusion should be interpreted with care. Future research should examine the optimal intensity and length of interventions.

Limitations

This meta-analysis has a number of limitations. First, because the risk of bias in the included studies was high, these results must be interpreted with caution. Second, the overall number of studies for the follow-up points and for some subgroups is small, limiting the strength of conclusions that can be drawn from these results. Third, we found a moderate degree of heterogeneity for the primary outcome that was reduced when analyzing the level of risk of bias and guidance subgroups; nevertheless, the number of comparisons in some subgroups was small and did occasionally overlap concerning individual studies. Fourth, we are mindful of the possibility that despite our thorough literature search, we might have missed a relevant study. Finally, because of the lack and inconsistency of information provided by the included studies, we were unable to analyze the effect of potentially relevant moderators of the treatment effect, such as the effect of adherence to the intervention on the overall outcome.

Future Recommendations

In future studies, preregistration of trials, an adequate calculation of the sample size, a more detailed description of allocation concealment, and appropriate methods to account for missing values are strongly recommended. We observed that a growing number of studies adhered to the intention-to-treat principle (ie, by employing mixed-effects models); nevertheless, especially for those studies, we recommend to ensure that the reported descriptive statistics are based on adequate methods to handle missing values and do not solely rely on complete cases. This will ensure adherence to the intention-to-treat principle throughout all statistics.

Although in most studies the therapeutic approach that was used in the intervention (eg, CBT) is often well described, it would be desirable for future studies to also specify the theoretical model that was used to develop the intervention (eg, conservation of resources theory, transactional model of stress).

Overall, more methodologically rigorous studies with a low risk of bias are needed to assess the effect of, for example, particular characteristics of interventions, such as treatment latitude, different levels of guidance, and different types of interventions in a direct comparison. Moreover, information on the number of participants who adhered to the intervention and details on co-interventions alongside training would be insightful. Future research should also test Web- and computer-based stress management interventions against the highest standard in this field (ie, classical face-to-face stress management interventions) and should more frequently include longer follow-up periods (eg, up to 6 months). Data on the cost-effectiveness of such interventions would also be highly relevant.

Conclusions

Despite the limitations discussed, it appears safe to conclude that Web- and computer-based interventions can be effective. In particular, interventions that include guidance from an online coach, are of medium length, or that are based on (third-wave) CBT lead to moderate improvements in stress levels. Initial evidence also suggests that the effects can be maintained up to 6 months. Whereas research and practice on traditional face-to-face interventions have been prolific [8,9], research on the efficacy and dissemination of Web-based stress-management interventions is still at the beginning despite the high potential and reach of such interventions. This work draws attention to the need for further studies on the efficacy, cost-effectiveness, and mechanisms of change of such interventions. In summary, the integration of evidence-based Web-based stress management interventions into health care systems has the potential to make a valuable contribution to reducing stress-related mental health problems on a large scale.

Acknowledgments

We would like to acknowledge the European Union (EFRE) for funding this project within the Lueneburg Innovation Incubator, TM 1.1 (project number: CCI 2007DE161PR001). Furthermore, we would like to thank Stefanie Christmann for assisting in the coding of the studies. We would also like to thank Douglas W Billings, Filip Drozd, Katherine Eisen, Yves Hänggi, Jeroen Ruwaard, Benjamin Wiegand, Viann Nguyen-Feng, and Ruth Q Wolever for providing further information on their studies.

Authors' Contributions

EH, DE, DL, MB, PC, and HR conceptualized the study. EH, DE, DL, and PC developed the search string. EH, SN, and DE performed the search, selected the studies, and collected the data. EH, DE, and DL analyzed and interpreted the data. EH drafted the manuscript; DE supervised the writing process; and DL, PC, MB, SN, and HR critically revised it. All authors agreed on the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 10KB - jmir_v19i2e32_app1.pdf]

Multimedia Appendix 2

Selected intervention characteristics of included studies.

[PDF File (Adobe PDF File), 119KB - jmir_v19i2e32_app2.pdf]

References

1. Hammen C. Stress and depression. Annu Rev Clin Psychol 2005 Apr;1:293-319. [doi: 10.1146/annurev.clinpsy.1.102803.143938] [Medline: 17716090]
2. Akerstedt T. Psychosocial stress and impaired sleep. Scand J Work Environ Health 2006 Dec;32(6):493-501 [FREE Full text] [Medline: 17173205]
3. Kraatz S, Lang J, Kraus T, Münster E, Ochsmann E. The incremental effect of psychosocial workplace factors on the development of neck and shoulder disorders: a systematic review of longitudinal studies. Int Arch Occup Environ Health 2013 May;86(4):375-395. [doi: 10.1007/s00420-013-0848-y] [Medline: 23549669]
4. Richardson S, Shaffer JA, Falzon L, Krupka D, Davidson KW, Edmondson D. Meta-analysis of perceived stress and its association with incident coronary heart disease. Am J Cardiol 2012 Dec 15;110(12):1711-1716 [FREE Full text] [doi: 10.1016/j.amjcard.2012.08.004] [Medline: 22975465]
5. Steptoe A, Kivimäki M. Stress and cardiovascular disease. Nat Rev Cardiol 2012 Jun;9(6):360-370. [doi: 10.1038/nrcardio.2012.45] [Medline: 22473079]
6. Hassard J, Teoh K, Cox T, Dewe P, Cosmar M, Grundler R, et al. Calculating the Cost of Work-Related Stress and Psychosocial Risks. Luxembourg: Publications Office of the European Union; 2014.
7. 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]
8. van der Klink JJ, Blonk RW, Schene AH, van Dijk FJ. The benefits of interventions for work-related stress. Am J Public Health 2001 Feb;91(2):270-276. [Medline: 11211637]
9. 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]
10. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-869; quiz 870. [doi: 10.1586/ern.12.63] [Medline: 22853793]
11. 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]
12. 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]
13. Lin J, Ebert DD, Lehr D, Berking M, Baumeister H. [Internet based cognitive behavioral interventions: state of the art and implementation possibilities in rehabilitation]. Rehabilitation (Stuttg) 2013 Jun;52(3):155-163. [doi: 10.1055/s-0033-1343491] [Medline: 23761204]
14. Ebert DD, Zarski AC, 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]
15. van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CFIII, Beekman AT, et al. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. Int J Epidemiol 2014 Apr;43(2):318-329 [FREE Full text] [doi: 10.1093/ije/dyt175] [Medline: 24760873]
16. Buntrock C, Ebert DD, Lehr D, Cuijpers P, Riper H, Smit F, et al. Evaluating the efficacy and cost-effectiveness of web-based indicated prevention of major depression: design of a randomised controlled trial. BMC Psychiatry 2014;14:25 [FREE Full text] [doi: 10.1186/1471-244X-14-25] [Medline: 24485283]
17. Buntrock C, Ebert DD, Lehr D, Smit F, Riper H, Berking M, et al. Effect of a Web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: a randomized clinical trial. JAMA 2016 May 3;315(17):1854-1863. [doi: 10.1001/jama.2016.4326] [Medline: 27139058]
18. Ebert DD, Buntrock C, Cuijpers P. Online intervention for prevention of major depression--reply. JAMA 2016;316(8):881-882. [doi: 10.1001/jama.2016.9586] [Medline: 27552628]
19. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. Cogn Behav Ther 2009;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
20. Cheng SK, Dizon J. Computerised cognitive behavioural therapy for insomnia: a systematic review and meta-analysis. Psychother Psychosom 2012;81(4):206-216. [doi: 10.1159/000335379] [Medline: 22585048]
21. Billings DW, Cook RF, Hendrickson A, Dove DC. A web-based approach to managing stress and mood disorders in the workforce. J Occup Environ Med 2008 Aug;50(8):960-968. [doi: 10.1097/JOM.0b013e31816c435b] [Medline: 18695455]
22. Hänggi Y. Kann web-basierte Stressprävention erfolgreich sein? [Can web-based parental training on coping with familiy stress be successful?]. Zeitschrift für Klinische Psychologie und Psychotherapie 2006 Jul 1;35(3):169-177. [doi: 10.1026/1616-3443.35.3.169]
23. Rose RD, Buckey JC, Zbozinek TD, Motivala SJ, Glenn DE, Cartreine JA, et al. A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. Behav Res Ther 2013 Feb;51(2):106-112. [doi: 10.1016/j.brat.2012.11.003] [Medline: 23262118]

24. Ruwaard J, Lange A, Bouwman M, Broeksteeg J, Schrieken B. E-mailed standardized cognitive behavioural treatment of work-related stress: a randomized controlled trial. Cogn Behav Ther 2007;36(3):179-192. [doi: 10.1080/16506070701381863] [Medline: 17852171]
25. Umanodan R, Shimazu A, Minami M, Kawakami N. Effects of computer-based stress management training on psychological well-being and work performance in japanese employees: a cluster randomized controlled trial. Ind Health 2014;52(6):480-491 [FREE Full text] [Medline: 25055847]
26. Zetterqvist K, Maanmies J, Ström L, Andersson G. Randomized controlled trial of internet-based stress management. Cogn Behav Ther 2003;32(3):151-160. [doi: 10.1080/16506070302316] [Medline: 16291546]
27. Aikens KA, Astin J, Pelletier KR, Levanovich K, Baase CM, Park YY, et al. Mindfulness goes to work: impact of an online workplace intervention. J Occup Environ Med 2014 Jul;56(7):721-731. [doi: 10.1097/JOM.0000000000000209] [Medline: 24988100]
28. Allexandre D, Bernstein AM, Walker E, Hunter J, Roizen MF, Morledge TJ. A Web-based mindfulness stress management program in a corporate call center: a randomized clinical trial to evaluate the added benefit of onsite group support. J Occup Environ Med 2016 Mar;58(3):254-264 [FREE Full text] [doi: 10.1097/JOM.0000000000000680] [Medline: 26949875]
29. Cavanagh K, Strauss C, Cicconi F, Griffiths N, Wyper A, Jones F. A randomised controlled trial of a brief online mindfulness-based intervention. Behav Res Ther 2013 Sep;51(9):573-578. [doi: 10.1016/j.brat.2013.06.003] [Medline: 23872699]
30. Drozd F, Raeder S, Kraft P, Bjørkli CA. Multilevel growth curve analyses of treatment effects of a Web-based intervention for stress reduction: randomized controlled trial. J Med Internet Res 2013;15(4):e84 [FREE Full text] [doi: 10.2196/jmir.2570] [Medline: 23607962]
31. Ebert DD, Heber E, Berking M, Riper H, Cuijpers P, Funk B, et al. Self-guided internet-based and mobile-based stress management for employees: results of a randomised controlled trial. Occup Environ Med 2016 May;73(5):315-323. [doi: 10.1136/oemed-2015-103269] [Medline: 26884049]
32. Ebert DD, Lehr D, Heber E, Riper H, Cuijpers P, Berking M. Internet- and mobile-based stress management for employees with adherence-focused guidance: efficacy and mechanism of change. Scand J Work Environ Health 2016 Sep 01;42(5):382-394. [doi: 10.5271/sjweh.3573] [Medline: 27249161]
33. Glück TM, Maercker A. A randomized controlled pilot study of a brief web-based mindfulness training. BMC Psychiatry 2011;11:175 [FREE Full text] [doi: 10.1186/1471-244X-11-175] [Medline: 22067058]
34. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
35. Ly KH, Asplund K, Andersson G. Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. Internet Interventions 2014 Jul;1(3):95-101. [doi: 10.1016/j.invent.2014.06.003]
36. Morledge TJ, Allexandre D, Fox E, Fu AZ, Higashi MK, Kruzikas DT, et al. Feasibility of an online mindfulness program for stress management-a randomized, controlled trial. Ann Behav Med 2013 Oct;46(2):137-148 [FREE Full text] [doi: 10.1007/s12160-013-9490-x] [Medline: 23632913]
37. Wolever RQ, Bobinet KJ, McCabe K, Mackenzie ER, Fekete E, Kusnick CA, et al. Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. J Occup Health Psychol 2012 Apr;17(2):246-258. [doi: 10.1037/a0027278] [Medline: 22352291]
38. Wiegand B, Luedtke K, Friscia D, Nair M, Aleles M, McCloskey R. Efficacy of a comprehensive program for reducing stress in women: a prospective, randomized trial. Curr Med Res Opin 2010 Apr;26(4):991-1002. [doi: 10.1185/03007991003688193] [Medline: 20201624]
39. Hinman M, Ezzo L, Hunt D, Mays A. Computerized exercise program does not affect stress levels of asymptomatic VDT users. J Occup Rehab 1997 Mar;7(1):45-51. [doi: 10.1007/BF02765875]
40. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [Medline: 22008217]
41. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to Meta-Analysis. Chichester, UK: John Wiley & Sons; 2009.
42. Hintz S, Frazier PA, Meredith L. Evaluating an online stress management intervention for college students. J Couns Psychol 2015 Apr;62(2):137-147. [doi: 10.1037/cou0000014] [Medline: 24635586]
43. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]
44. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biol Psychiatry 2006 Jun 1;59(11):990-996. [doi: 10.1016/j.biopsych.2005.09.014] [Medline: 16368078]
45. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 1988 Jun 30;318(26):1728-1733. [doi: 10.1056/NEJM198806303182605] [Medline: 3374545]
46. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]

47. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. BMJ 2007 Nov 3;335(7626):914-916 [FREE Full text] [doi: 10.1136/bmj.39343.408449.80] [Medline: 17974687]
48. Orsini N, Higgins J, Bottai M, Buchan I. EconPapers. 2005 Feb 16. Heterogi: Stata module to quantify heterogeneity in a meta-analysis URL: http://econpapers.repec.org/RePEc:boc:bocode:s449201 [accessed 2017-01-25] [WebCite Cache ID 6nmdxJgEp]
49. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [Medline: 9310563]
50. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000 Jun;56(2):455-463. [Medline: 10877304]
51. Lange A, van de Ven J, Schrieken B, Smit M. ‘Interapy’ Burn-out: Prävention und Behandlung von Burn-out über das Internet [‘Interapy’ burnout: prevention and therapy of burnout via the Internet]. Verhaltenstherapie 2004 Oct 4;14(3):190-199. [doi: 10.1159/000080915]
52. Chiauzzi E, Brevard J, Thum C, Thurn C, Decembrele S, Lord S. MyStudentBody-Stress: an online stress management intervention for college students. J Health Commun 2008 Sep;13(6):555-572. [doi: 10.1080/10810730802281668] [Medline: 18726812]
53. Eisen KP, Allen GJ, Bollash M, Pescatello LS. Stress management in the workplace: a comparison of a computer-based and an in-person stress-management intervention. Comput Hum Behav 2008;24:486-496. [doi: 10.1016/j.chb.2007.02.003]
54. Hasson H, Anderberg UM, Theorell T, Arnetz BB. Psychophysiological effects of a web-based stress management system: a prospective, randomized controlled intervention study of IT and media workers [ISRCTN54254861]. BMC Public Health 2005 Jul 25;5:78 [FREE Full text] [doi: 10.1186/1471-2458-5-78] [Medline: 16042796]
55. Nguyen-Feng VN, Frazier PA, Greer CS, Howard KG, Paulsen JA, Meredith L, et al. A randomized controlled trial of a web-based intervention to reduce distress among students with a history of interpersonal violence. Psychol Violen 2015;5(4):444-454. [doi: 10.1037/a0039596]
56. Frazier P, Meredith L, Greer C, Paulsen JA, Howard K, Dietz LR, et al. Randomized controlled trial evaluating the effectiveness of a web-based stress management program among community college students. Anxiety Stress Copin 2015;28(5):576-586. [doi: 10.1080/10615806.2014.987666] [Medline: 25420030]
57. Yamagishi M, Kobayashi T, Nakamura Y. Effects of web-based career identity training for stress management among Japanese nurses: a randomized control trial. J Occup Health 2008;50(2):191-193 [FREE Full text] [Medline: 18403870]
58. Johnson BT, Scott-Sheldon LA, Carey MP. Meta-synthesis of health behavior change meta-analyses. Am J Public Health 2010 Nov;100(11):2193-2198 [FREE Full text] [doi: 10.2105/AJPH.2008.155200] [Medline: 20167901]
59. Junge MN, Lehr D, Bockting CL, Berking M, Riper H, Cuijpers P, et al. For whom are internet-based occupational mental health interventions effective? Moderators of internet-based problem-solving training outcome. Internet Interventions 2015 Mar;2(1):39-47. [doi: 10.1016/j.invent.2014.11.007]
60. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
61. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
62. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
63. Zarski AC, Lehr D, Berking M, Riper H, Cuijpers P, Ebert DD. Adherence to Internet-based mobile-supported stress management: a pooled analysis of individual participant data from three randomized controlled trials. J Med Internet Res 2016;18(6):e146 [FREE Full text] [doi: 10.2196/jmir.4493] [Medline: 27357528]
64. Ebert DD, Lehr D, Smit F, Zarski AC, Riper H, Heber E, et al. Efficacy and cost-effectiveness of minimal guided and unguided internet-based mobile supported stress-management in employees with occupational stress: a three-armed randomised controlled trial. BMC Public Health 2014;14:807 [FREE Full text] [doi: 10.1186/1471-2458-14-807] [Medline: 25099533]
65. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. A meta-analysis. J Psychosom Res 2004 Jul;57(1):35-43. [doi: 10.1016/S0022-3999(03)00573-7] [Medline: 15256293]
66. Vøllestad J, Nielsen MB, Nielsen GH. Mindfulness- and acceptance-based interventions for anxiety disorders: a systematic review and meta-analysis. Br J Clin Psychol 2012 Sep;51(3):239-260. [doi: 10.1111/j.2044-8260.2011.02024.x] [Medline: 22803933]
67. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
68. Hunot V, Moore TH, Caldwell DM, Furukawa TA, Davies P, Jones H, et al. 'Third wave' cognitive and behavioural therapies versus other psychological therapies for depression. Cochrane Database Syst Rev 2013;10:CD008704. [doi: 10.1002/14651858.CD008704.pub2] [Medline: 24142844]
69. Cuijpers P, Huibers M, Ebert DD, Koole SL, Andersson G. How much psychotherapy is needed to treat depression? A metaregression analysis. J Affect Disord 2013 Jul;149(1-3):1-13. [doi: 10.1016/j.jad.2013.02.030] [Medline: 23528438]

Abbreviations

CBT: cognitive behavioral therapy
NNT: number needed to treat
RCT: randomized controlled trial
TWC: third-wave cognitive behavioral therapy

Edited by G Eysenbach; submitted 23.03.16; peer-reviewed by J García Campayo, C Buntrock; comments to author 20.04.16; revised version received 15.11.16; accepted 23.11.16; published 17.02.17.

Please cite as:
Heber E, Ebert DD, Lehr D, Cuijpers P, Berking M, Nobis S, Riper H
The Benefit of Web- and Computer-Based Interventions for Stress: A Systematic Review and Meta-Analysis
J Med Internet Res 2017;19(2):e32
URL: http://www.jmir.org/2017/2/e32/
doi:10.2196/jmir.5774
PMID:28213341

Original Paper

A Mobile Device App to Reduce Time to Drug Delivery and Medication Errors During Simulated Pediatric Cardiopulmonary Resuscitation: A Randomized Controlled Trial

Johan N Siebert[1*], MD; Frederic Ehrler[2*], PhD; Christophe Combescure[3], PhD; Laurence Lacroix[1], MD; Kevin Haddad[1], RN; Oliver Sanchez[4], MD; Alain Gervaix[1], MD; Christian Lovis[2], MPH, MD; Sergio Manzano[1], MD

[1]Department of Pediatric Emergency Medicine, Geneva Children's Hospital, University Hospitals of Geneva, Geneva, Switzerland

[2]Department of Radiology and Medical Informatics, Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland

[3]Department of Health and Community Medicine, Division of Clinical Epidemiology, University of Geneva and University Hospitals of Geneva, Geneva, Switzerland

[4]Department of Pediatric Surgery, Geneva Children’s Hospital, University Hospitals of Geneva, Geneva, Switzerland

[*]these authors contributed equally

Corresponding Author:
Johan N Siebert, MD
Department of Pediatric Emergency Medicine
Geneva Children's Hospital
University Hospitals of Geneva
Avenue de la Roseraie, 47
Geneva, 1205
Switzerland
Phone: 41 795534072
Fax: 41 223725423
Email: Johan.Siebert@hcuge.ch

Abstract

Background: During pediatric cardiopulmonary resuscitation (CPR), vasoactive drug preparation for continuous infusion is both complex and time-consuming, placing children at higher risk than adults for medication errors. Following an evidence-based ergonomic-driven approach, we developed a mobile device app called Pediatric Accurate Medication in Emergency Situations (PedAMINES), intended to guide caregivers step-by-step from preparation to delivery of drugs requiring continuous infusion.

Objective: The aim of our study was to determine whether the use of PedAMINES reduces drug preparation time (TDP) and time to delivery (TDD; primary outcome), as well as medication errors (secondary outcomes) when compared with conventional preparation methods.

Methods: The study was a randomized controlled crossover trial with 2 parallel groups comparing PedAMINES with a conventional and internationally used drugs infusion rate table in the preparation of continuous drug infusion. We used a simulation-based pediatric CPR cardiac arrest scenario with a high-fidelity manikin in the shock room of a tertiary care pediatric emergency department. After epinephrine-induced return of spontaneous circulation, pediatric emergency nurses were first asked to prepare a continuous infusion of dopamine, using either PedAMINES (intervention group) or the infusion table (control group), and second, a continuous infusion of norepinephrine by crossing the procedure. The primary outcome was the elapsed time in seconds, in each allocation group, from the oral prescription by the physician to TDD by the nurse. TDD included TDP. The secondary outcome was the medication dosage error rate during the sequence from drug preparation to drug injection.

Results: A total of 20 nurses were randomized into 2 groups. During the first study period, mean TDP while using PedAMINES and conventional preparation methods was 128.1 s (95% CI 102-154) and 308.1 s (95% CI 216-400), respectively (180 s reduction, *P*=.002). Mean TDD was 214 s (95% CI 171-256) and 391 s (95% CI 298-483), respectively (177.3 s reduction, *P*=.002). Medication errors were reduced from 70% to 0% (*P*<.001) by using PedAMINES when compared with conventional methods.

Conclusions: In this simulation-based study, PedAMINES dramatically reduced TDP, to delivery and the rate of medication errors.

(J Med Internet Res 2017;19(2):e31) doi:10.2196/jmir.7005

KEYWORDS

resuscitation; medication errors; pharmaceutical preparations; pediatrics; biomedical technology

Introduction

Immediate (level 1) triage represents 175,000 patient visits every year in US pediatric emergency departments (PED) [1]. Among them, 6700 to 15,000 cases are due to out-of-hospital cardiac arrest (OHCA) [2-4], including 6000 related to nontraumatic origins [5], and 5800 to 10,000 to in-hospital cardiac arrest (INHCA) [6,7]. In our institution in 2014, cardiopulmonary resuscitation (CPR) accounted for 0.5% of almost 28,000 pediatric visits (0-15 years old). In CPR, time is a decisive success criterion. During the first 15 min, survival and favorable neurological outcome decrease linearly by 2.1% and 1.2% per minute, respectively [8], and are negatively affected by drug preparation (TDP) and delivery time (TDD) [9]. In a study with adults in cardiac arrest, the chance of return of spontaneous circulation (ROSC) was decreased by 4% for every 1-min delay in delivery of vasopressor [10].

Prolonged resuscitation time may result from TDP [11]. During some critical situations such as postcardiac arrest ROSC or septic shock, preparing intravenous (IV) vasoactive drugs for continuous infusion is particularly challenging. Quickly, accurately, and safely preparing and administering drugs in a stressful environment is complex and time-consuming [12-14]. The need for individual specific weight-based drug dose calculation and preparation and a lower dosing-error tolerance [15] place children at higher risk than adults for errors [16-18] and may result in life-threatening outcomes. Medication errors have been reported in up to 41% of pediatric resuscitations, the most common being incorrect medication dosage found in up to 65% of cases [19]. Proper preparation and delivery of these drugs could favorably affect the pediatric CPR outcomes.

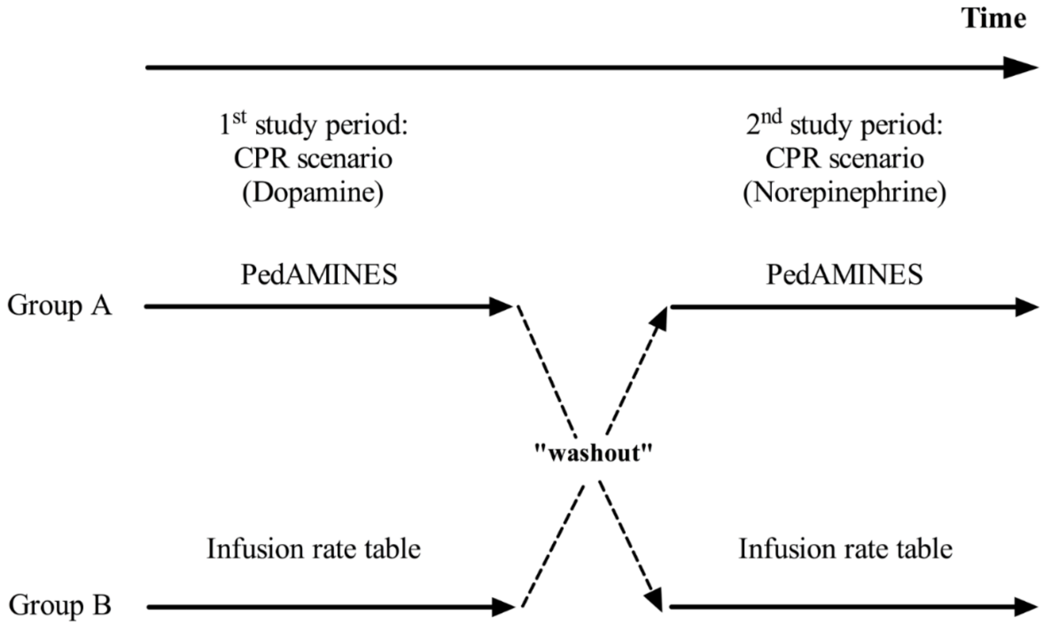
To address these problems, we followed a cognitive and evidence-based ergonomic driven approach [20] to develop an innovative and customizable tablet app, called Pediatric Accurate Medication in Emergency Situations (PedAMINES). This app was designed to support nurses and physicians step-by-step from order to delivery of a wide range of drugs in real time, including those requiring continuous infusion [21]. The development of the app was followed by a study aiming to assess its impact on the error rate and time needed from drug prescription to administration. We hypothesized that PedAMINES would first reduce the TDP and TDD, and second, reduce medication errors during pediatric CPRs when compared with conventional preparation methods.

Methods

Study Design

The study was a prospective, randomized controlled crossover trial with 2 parallel groups (Figure 1) comparing PedAMINES [21] with a conventional and internationally used drugs infusion rate table method (Frank Shann conventional drug infusion rate table [22]; Multimedia Appendix 1) in the preparation of continuous drug infusion, during a standardized simulation-based pediatric postcardiac arrest scenario. No changes were made on the app or on the intervention during the study.

Figure 1. Study design: A two-period, randomized controlled, two-arm crossover study.



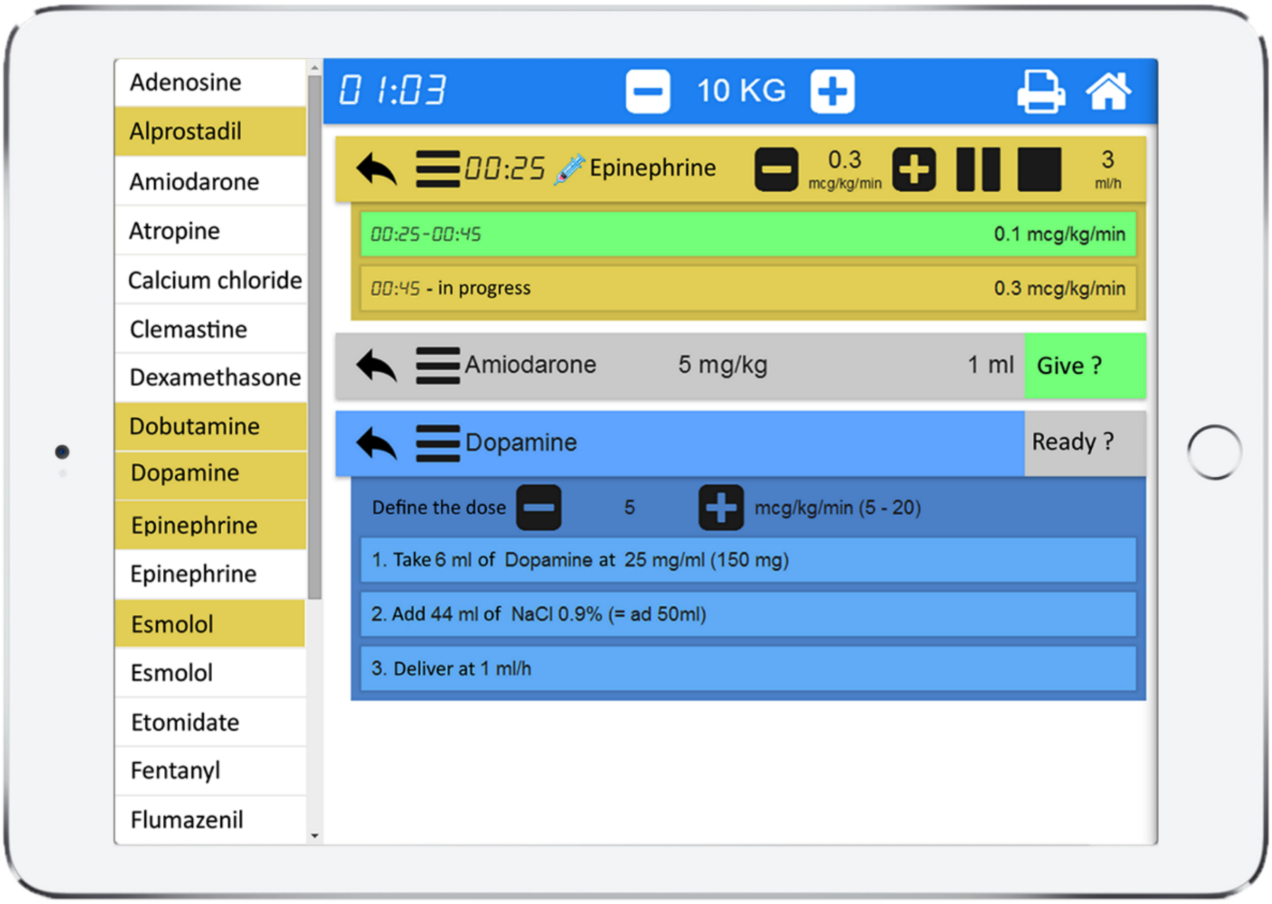
Selection of Participants

Certified pediatric emergency nurses were eligible if they were actively practicing in our PED. Shift-working nurses were randomly recruited on the day of the study by a blinded, noninvestigator, person on a random list. Written informed consent was obtained from all the participants before their voluntary involvement. Study participants were neither involved in the design of the app, nor in the study design, choice of outcome measures, or the study conduct. A senior specialized nurse in pediatric emergency medicine, simulation, and teaching, being an investigator of the study, has participated in the app and study design. No participants were asked to advice on interpretation or writing up of results. Results of the study were disseminated to the participants after the completion of the study.

Setting

The study was conducted in a PED of a tertiary hospital with approximately 28,000 visits per year. The PedAMINES app lists all the available resuscitation drugs, with doses automatically adapted to the weight or age of the patient. Evidence-based development of tools is an efficient way to develop apps that support clinicians [20]. The development of PedAMINES followed a user-centered approach with emergency department (ED) caregivers, as well as software developers and ergonomists. This team worked tightly together and the app development was mainly based on CPR observations and focus groups [21]. In this study, 6 drugs for continuous infusion and 19 drugs for direct IV injection were listed in the PedAMINES app and at the nurse's disposal. The list can be expanded and customized according to users' desires and to local drugs habits. By a simple touch, any of the listed drugs can be selected and its preparation detailed according to a standardized and simplified path. In the case of a continuous infusion, this path is composed of 3 steps: (1) drug selection, (2) dilution of the initial drug concentration, and (3) conversion of the prescribed dose rate in µg/kg/min into infusion pump rate in ml/h. For each drug, the exact amount to prepare is clearly displayed and thus avoids the necessity for calculations (Figure 2). This is based on the app's ability to automatically calculate the optimal weight-based final infusion pump rate and to describe the preparation sequence required to achieve it, independently of the nurse competency in this domain. The nurse may, at any time, interact with the app. The user can start, pause, stop, increase, or diminish the perfusion rate. Multiple drugs can be prepared and run in parallel. All actions performed by the nurses were sequentially saved locally on the device in historic files to preserve information that can be retrieved at any time for debriefing or medicolegal purposes. Historic files can also be erased or safely exported and saved on the institutional electronic health record.

Figure 2. PedAMINES screenshot. List of bolus IV drugs (white boxes) and drugs for continuous infusion (yellow boxes) are selectable in the left margin of the application. The right window shows drugs selected by the nurse for a ten kilograms child. In this screenshot example, epinephrine is being delivered at an infusion rate of 0.3 mcg/kg/min. Amiodarone is selected and ready to be injected, waiting for nurse's confirmation. Dopamine is being prepared by the nurse following a descriptive and detailed path automatically calculated by the application. The printer logo in the upper right corner indicates that all actions performed by the nurses are sequentially saved in historic files that can be retrieved and printed at any time.



00:25 Epinephrine 0.3 mcg/kg/min 3 ml/h

00:25-00:45 0.1 mcg/kg/min

00:45 - in progress 0.3 mcg/kg/min

Amiodarone 5 mg/kg 1 ml Give ?

Dopamine Ready ?

Define the dose 5 mcg/kg/min (5 - 20)

1. Take 6 ml of Dopamine at 25 mg/ml (150 mg)
2. Add 44 ml of NaCl 0.9% (= ad 50ml)
3. Deliver at 1 ml/h

Intervention and Resuscitation Scenario

We created a standardized simulation scenario on a high-fidelity manikin (Laerdal SimBaby) in the pediatric shock room of our PED. Consistent with standard emergency medicine practice, we created resuscitation teams comprising the study-participating nurse, an emergency pediatrician leading the resuscitation and being a Pediatric Advanced Life Support (PALS) instructor, and a second nurse (both part of the investigators team) to assist with resuscitation by performing chest compressions and bag-valve-mask ventilation according to the pediatrician instructions. A certified technician operated the simulator. Except for the participating nurse, the members of the resuscitation team were unchanged across all the scenarios and were the investigators of the study.

On the day of participation, nurses were given a survey collecting data regarding their demographics, nursing, and computer experience. After random allocation, each participating nurse received a standardized 5-min training session on how to use the app PedAMINES to familiarize them uniformly with it. Then, the nurses were asked to perform a 15-min highly realistic CPR scenario, including postreturn of spontaneous circulation (ROSC). It was standardized to follow the same chronological progression and provided on the same manikin. The scenario was conducted in situ in our shock room to increase realism. When entering the shock room, the nurses were asked to assist the physician to perform a 2-min full course massage and ventilation (15:2 ratio) maneuver for a child with asystole to increase their own stress level. On the basis of the American Heart Association (AHA), pediatric cardiac arrest algorithm for asystole, a bolus of 0.01 mg/kg epinephrine (0.1 mL/kg of 1:10,000 concentration) was administered. ROSC ensued with hypotension. At that time, a clinical statement to recognize the life-threatening condition of the patient, including his weight and age, was given to the nurses. The nurses were then asked to prepare a 5 µg/kg/min continuous infusion of dopamine for a 7-kg boy either with the help of PedAMINES first (allocation group A) or following the conventional method first (ie, Shann infusion rate table [22], group B; Multimedia Appendix 1). All participants had equal experience and competence with the Shann method. At the end of the instruction, the timed scenario began. To ensure that participants heard and understood the order correctly, they had to confirm it verbally, and written transcriptions were checked and video-recorded. When the drug was ready to be injected, the nurse was asked to deliver it to the patient using a syringe pump already in place. The nurse was then asked to perform a 1-min "washout" distraction maneuver by aspirating secretions in the throat of the manikin. At this moment, the crossover occurred. The nurse was asked to prepare a 0.1 µg/kg/min continuous infusion of norepinephrine by crossing the procedure (ie, group A allocation having used PedAMINES before was asked to use the conventional method, whereas group B allocation having used the conventional method before was asked to use PedAMINES; Figure 1). At the end of the order, the nurse was asked to prepare the drug. Given that norepinephrine preparation could present a challenge by requiring a more complex decimal-point dependent calculation with the Shann method, we rendered the tasks uniform by ordering a decimal-point final volume calculation for both Frank Shann and PedAMINES preparation methods. When the drug was ready to be injected, the nurse was asked to deliver it to the patient using a second syringe pump already in place. The delivery of both drugs required programming the same pump in a similar manner among all participants. Time elapsed after drug preparation until its delivery, that is, time needed to set up the pump, was assessed for all participants to ensure uniformity among participants. The measured deviation between the amount of drug delivered and the actual prescribed dose was given by the amount of drug in the syringe. This was verified by an examiner during the scenario and video-recorded. The beginning of the injection corresponded to the end of the scenario. At that time, the nurse had to recall and describe precisely how she prepared both the drugs.

Outcome Measures

The primary outcome was the elapsed time in seconds, in each allocation group, between the oral prescription by the physician and TDD by the nurse. TDD included TDP completion by the nurse. The secondary outcome was the medication dosage error rate in each allocation group, during the sequence from drug preparation to drug injection. Regarding both outcomes, we considered a 15 s difference in delivery of resuscitation drugs [23], and a 30% difference in the rate of medication errors to be clinically significant and sufficient to modify practice. At the end of the scenario, a 2 question questionnaire using a 10-point Likert scale (scored from 1 to 10 to avoid neutral answers) was submitted to participants. The questionnaire measured (1) the overall stress perceived (the question was "On a scale of 1 to 10, how much stress did you feel during the whole resuscitation scenario?") and (2) the satisfaction about the preparation method used during the resuscitation scenario (the question was "On a scale of 1 to 10, how much satisfaction did you get during the resuscitation scenario with the help of PedAMINES, and with the help of the infusion rate table?").

Methods of Measurement and Data Collection

Data collected during the scenario included (1) TDP, (2) TDD, and (3) final delivered drug concentration in µg/ml and infusion rate in µg/kg/min. All the actions (ie, primary and secondary outcomes) performed by the nurses during the scenario were automatically recorded and stored by the responsive simulator detectors and by several video cameras. To avoid assessment bias, 2 evaluators then independently reviewed these video recordings. In case of disagreement, a third independent evaluator helped reach a consensus. All actions performed with PedAMINES were automatically saved locally in log files for further analysis. Data were manually retrieved and entered into a Microsoft Excel spreadsheet (version 2011). The statistical software GraphPad Prism version 6.0h (GraphPad Software, Inc) was used for graph figures. Stata/IC version 14 (StataCorp) was used for descriptive analyses, and R version 2.15.2 (R Foundation for statistical computing) was used for statistical tests and 95% CI.

Sample Size

The primary objective of this study was to detect a difference in TDD of vasoactive drugs between the 2 groups. The sample size was calculated to detect a 15 s decrease in TDD between

2 independent groups with a power of 90% and a 2-sided risk alpha of .05. A previous study with pediatric emergency nurses has shown a median TDD of 69 s for the first dose of vasoactive drug to be given as direct IV infusion [23]. Assuming a standard deviation (SD) of 9 s for TDD in each group (based on a similar SD of 10 s estimated by Moreira et al [23]), 9 participants per group were required. To prevent a potential loss of power due to misspecification of assumptions, it was necessary to recruit 10 nurses per group (total sample size: 20 nurses). In case of a carryover effect, this sample size calculation was sufficient to evaluate PedAMINES's effect within the first period of the trial.

Randomization and Blinding

We randomly assigned nurses in a 1:1 ratio with a Web-based software [24]. Blinding to the purpose of the study during recruitment was maintained to minimize preparation bias. Nurses were unblinded after randomization. Allocation concealment was ensured with sealed envelopes and was not released until the nurses started the scenario.

Statistical Analysis

Primary Outcome

For TDD (and TDP), the mean times were reported with 95% CI for each arm and each study period to investigate a potential carryover effect. As a carryover effect was suspected, intervention arms were compared within each study period using *t* tests for independent groups. No paired data were compared.

Secondary Outcomes

The rate of medication errors was the proportion of nurses making a preparation error. The rate of medication errors was reported with 95% CI (Clopper-Pearson method) with each method and by study period to investigate a potential carryover effect. The error rates for each method were globally assessed and compared using McNemar test as observations were paired. Differences in error rates were reported with 95% CIs.

Errors were also measured as the deviation in percent from the amount of delivered drug compared with the original dose prescribed by the physician. Absolute deviation was analyzed. The mean (SD) difference in deviation obtained with each method was reported with 95% CI. A *t* test for paired data was used to compare interventions. Mean differences were also reported by randomized group and by crossover period. Means and SDs were determined for stress and satisfaction scores of individuals for each questionnaire item and reported with descriptive statistics.

Ethics and Informed Consent

The study was approved by the institutional ethics committee and a trial registration number was not required. If we regard simulation as a translational science with outcomes categorized as T1 (results achieved in the simulated setting), T2 (improved health care delivery in the real clinical setting), and T3 (improved patient outcomes), our trial was a T-1 level study and, as for many other simulation studies, did not require registration according to ICMJE definition. Written informed consent was obtained from all participants before their voluntary involvement. It was conducted in accordance with the principles of the Declaration of Helsinki, the standards of Good Clinical Practice, and Swiss regulatory requirements.

Results

Study Participants

In June 2015, 20 certified pediatric emergency nurses participated and completed the study with no dropout (Figure 3). The demographic results are summarized in Table 1.

Table 1. Participants' demographic and clinical characteristics.

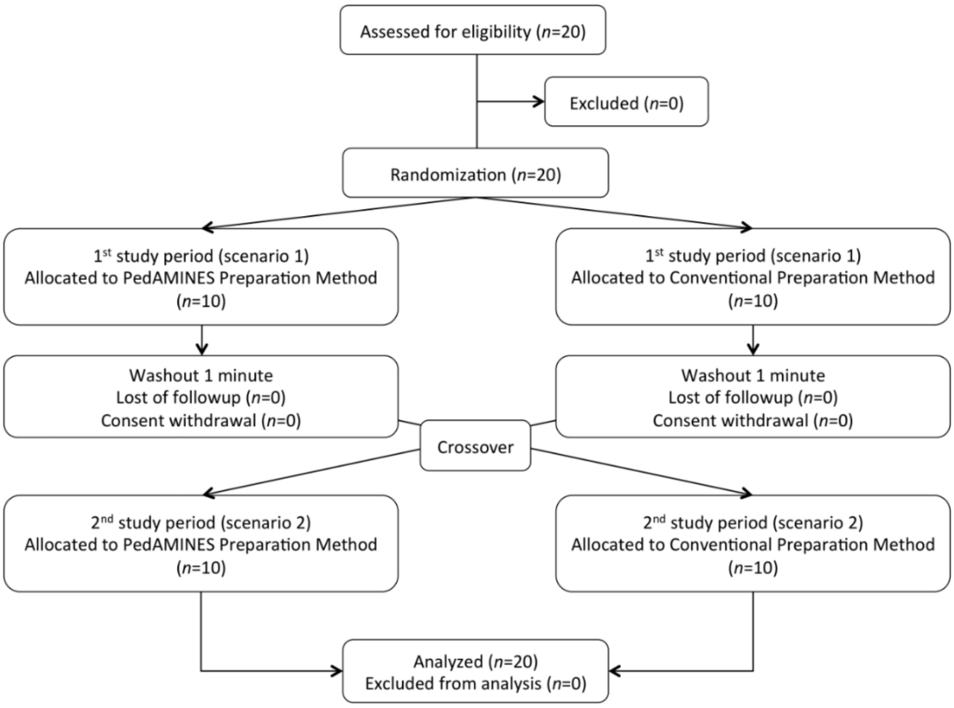
Demographics and clinical characteristics	Randomization arm (first study period)	
	PedAMINES[a] (n=10)	Infusion rate table (n=10)
Age (years), mean (SD)	42.9 (6.4)	42.4 (10.5)
Sex (female), n (%)	8 (80)	7 (70)
Number of years since nurse certification, mean (SD)	19.2 (7.7)	18.5 (11.7)
Number of years since pediatric ED[b] certification, mean (SD)	9.9 (6.4)	9.8 (6.5)
Own and use a tab or mobile phone at home, n (%)	9 (90)	9 (90)
CPRs[c] having required vasoactive drugs Preparation for continuous infusion in the past 3 years, median (interquartile range)	2.0 (0.0-2.30)	1.0 (1.0-2.0)
Simulated CPR scenarios in the past 3 years, median (interquartile range)	7.0 (3.8-12.0)	5.5 (3.8-8.3)

[a] PedAMINES: Pediatric Accurate Medication in Emergency Situations.

[b] ED: emergency department.

[c] CPR: cardiopulmonary resuscitation.

Figure 3. Patient flowchart for randomized controlled trial on preparation of continuous drug infusion by nurses in simulation-based pediatric cardio-pulmonary resuscitation cardiac arrest scenario.



TDP and TDD

When using PedAMINES, the mean time either from drug prescription to preparation or to delivery was nearly equivalent in both study periods (128.1 s vs 143.7 s or 213.5 s vs 221.8 s, *P*=.71; Table 2). Using the conventional method, the mean time from prescription to preparation or to delivery was lower in the second period than in the first period (308.1 s vs 198.4 s, *P*=.03 and 390.8 s vs 276.7 s, *P*=.03, respectively; Table 2). These findings raised the suspicion of a carryover effect and comparisons between interventions were conducted for each period separately. In the first study period, the TDP was decreased by 180.0 s (95% CI 86.5-273.5, *P*=.002) with PedAMINES and TDD by 177.3 s (95% CI 79.7-274.9, *P*=.002; Figure 4). These gains in time represented 58% and 45% of the mean time, respectively, when using the conventional methods. The variability of individual recorded times was lower with PedAMINES (TDP upper range value=222 s, TDD=320 s) than with the conventional method (TDP upper range value=545 s [8 out of 10 measures were higher than 222], TDD=657 s; Figure 4). In the second study period, the TDP was decreased by 54.7 s (95% CI 10.3-99.1, *P*=.02) with PedAMINES and TDD by 54.9 s (95% CI 1.6-108.2, *P*=.04; Figure 4).

Table 2. Mean time in seconds to drugs preparation and delivery.

Mean time	First study period (Dopamine)		Second study period (Norepinephrine)	
	PedAMINES[a] mean (95% CI)	Conventional method mean (95% CI)	PedAMINES mean (95% CI)	Conventional method mean (95% CI)
TDP[b]	128.1 (102.0-154.2)	308.1 (216.3-399.9)	143.7 (128.1-159.3)	198.4 (155.3-241.5)
SD	36.5	128.3	21.9	60.2
Time difference[c]	180.0 (86.5-273.5)		54.7 (10.3-99.1)	
TDD[d]	213.5 (170.6-256.4)	390.8 (298.3-483.3)	221.8 (198.2-245.4)	276.7 (226.2-327.2)
SD	59.9	129.3	33.0	70.6
Time difference[c,e]	177.3 (79.7-274.9)		54.9 (1.6-108.2)	

[a] PedAMINES: Pediatric Accurate Medication in Emergency Situations.

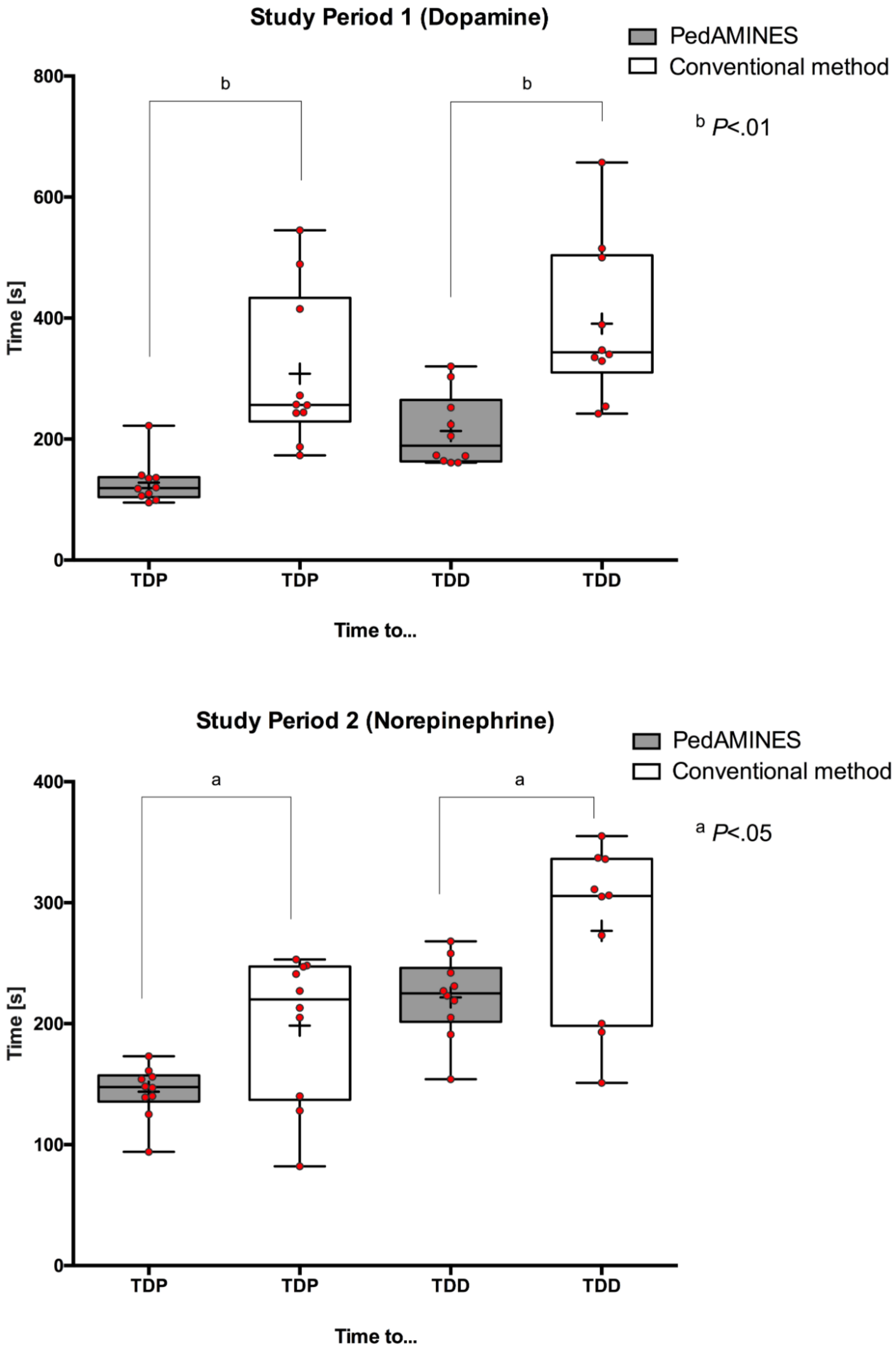
[b] TDP: time to drug preparation.

[c] Time difference represents time with the conventional method minus time with PedAMINES.

[d] TDD: time to drug delivery.

[e] See Multimedia Appendix 2 for TDP and TDD details for each nurse and for each drug, by study period.

Figure 4. Study period 1 (Dopamine) and 2 (Norepinephrine). Boxplots of elapsed time to drug preparation (TDP) and to drug delivery (TDD) in intervention group (PedAMINES) and control group (conventional method). Solid horizontal lines denote median and interquartile ranges; the whiskers go down to the smallest value and up to the largest; + denotes mean. Red open circles denote each individual value. Time is expressed in seconds.



Medication Error Rate

As data did not support a carryover effect in the medication error rate, both the study periods were pooled. Of the 20 drug doses delivered with PedAMINES, none (0%) was associated with medication errors. Of the 20 drug doses prepared with conventional methods, 14 were incorrect (70%; 95% CI

45.7-88.1, *P*<.001). Among the 8 errors committed during the first study period, 5 (63%, 5/8) were critical overdose errors ranging from 100% to 5233% (mean 1864%) of the normal prescribed dose (Table 3). Among the 6 errors committed during the second study period, 3 (50%, 3/6) ranged from 19% to 138% (mean 59%) of the normal prescribed dose and 2 (33%, 2/6) miscalculated preparations reached the right final dosage by chance (Table 3).

Table 3. Drug doses errors and deviation from the prescribed doses.

Outcomes	First study period (Dopamine)		Second study period (Norepinephrine)[a]	
	PedAMINES[b] n (%), 95% CI	Conventional method n (%), 95% CI	Conventional method n (%), 95% CI	PedAMINES n (%), 95% CI
All medication errors[c] (n=10)	0 (0), 0-31	8 (80), 44-98	6 (60), 26-88	0 (0), 0-31
Unpaired medication errors difference (n=10)	8 (80), 41.4-97.5		6 (60), 21.3-88.5	
Paired medication errors difference (N=20)	14 (70), 42.2-88.1			

[a] See Multimedia Appendix 3 for drug doses errors details for each nurse and for each drug by study period.

[b] PedAMINES: Pediatric Accurate Medication in Emergency Situations.

[c] Proportion of dosage with an error, where n denotes the number of drug dose errors that actually occurred and N denotes the total number of opportunities for errors to occur.

Questionnaire About Perceived Stress and Satisfaction

The questionnaire was completed and returned by 100% of the participants. Participants rated the overall perceived stress to be 7.1 (95% CI 6.1-8.1) on the Likert scale. They reported higher satisfaction when using PedAMINES for the preparation of drugs rather than conventional methods (9.3 [SD 1.2] vs 3.6 [SD 2.1], *P*<.001).

Discussion

Principal Findings

To our knowledge, this is the first study to investigate the benefit of a mobile app to improve delivery of continuous drug infusion during pediatric CPR. We found that TDP and TDD of vasoactive drugs for continuous infusion, as well as medication errors were dramatically reduced with the use of PedAMINES.

CPR is a continuum from the onset of resuscitation to immediate post-resuscitation care. Survival rate for CPR lasting less than 15 min is estimated at 41%, decreasing to 12% after 35 min [8]. A recent study reported that average time spent in the pediatric shock room was 46 min [25]. Care for postcardiac arrest patients is also time-sensitive. Postcardiac arrest syndrome (PCAS), characterized by brain injury, myocardial dysfunction, systemic ischemia or reperfusion response, and persistent precipitating disease, can develop precociously with poor outcome after ROSC [26,27]. Post-ROSC in-hospital mortality rates in children after nontraumatic OHCA or INHCA were estimated to be 70% [28] and 55% [29], respectively. If ROSC is quickly achieved and maintained after the onset of cardiac arrest, PCAS might be prevented [30]. Thus, early hemodynamic optimization improves the outcome of these patients [31,32]. The 2015 AHA guidelines recommended starting IV fluids and vasoactive drugs early in the postarrest phase, targeting a systolic blood pressure above the fifth percentile for age [33]. Commonly employed drugs for continuous infusion include dopamine (5-20 mcg/kg/min), norepinephrine, and epinephrine (both at 0.01-1 mcg/kg/min). These drugs should be available quickly. Moreira et al reported that the use of prefilled color-coded syringes during pediatric cardiac arrest can reduce the TDD of IV drugs by 27 s [23]. However, these results were limited to direct IV pushes requiring no prior preparation. Our work demonstrates that it is possible to drastically decrease TDD for IV drugs requiring complex upstream preparations and continuous infusion during the immediate postarrest phase. In our study, mean TDP, as well as mean TDD were significantly reduced by almost 180 s with PedAMINES. Even when considering the lower margins of errors of the CI, PedAMINES was still able to reduce TDP and TDD by approximately 1.5 min (4.5 min at the upper margins) and was largely higher than the minimal difference we set out to find when establishing the sample size. The interindividual upper time ranges were reduced by using the app, suggesting a worthwhile benefit if used in smaller hospitals where nurses and physicians are either rarely or not exposed to CPR, but have to use resuscitative drugs before patient transfer to a tertiary care center.

Although TDP and TDD were similar in the primary and secondary study periods when using PedAMINES, the gain provided by PedAMINES was less pronounced in the second study period, because TDP and TDD with the conventional method were lower among nurses with a PedAMINES experience than among nurses who had never used PedAMINES. The 1-min washout period may not have been long enough. Initially, we sought to estimate the effect of PedAMINES in the first study period. We wanted to determine whether changing habits with the use of our app by naïve nurses hitherto accustomed to a conventional preparation method resulted in a significant gain of time on the preparation of drugs compared with nurses using the conventional method. The second study period was intended to increase the overall power of our study. However, in this second period, the nurses could not be considered in exactly the same way as in the first period. Indeed, in the second period, we compared nurses who returned with a conventional preparation method after having used PedAMINES versus naïve nurses hitherto accustomed to a conventional preparation method that switched for PedAMINES for the first time. In any case, the intervention had a large impact in both periods. Given the carryover effect, the first study period was chosen as the most representative for the influence of PedAMINES on the primary outcome.

Medication errors are common in pediatric patients, accounting for 5-27% of all pediatric prescriptions and causing significant mortality and morbidity, including 7000 patient deaths each year in the United States [34]. One in every 32 prescriptions in a PED contains a 10-fold error on the recommended dose [17]. Kozer et al showed an error rate of 17% at the prescription level during pediatric resuscitations, with up to 10 times the recommended pediatric dose in more than 3% of these cases [35]. In the same study, 16% of the analyzed syringes showed a 20% dose deviation from the prescribed dose and up to 7% showed a deviation of more than 50%. Medication errors with infusions frequently result from mistake during preparation, due to wrong drug-volume calculations, imprecision with volume measurements or incorrect mixing during dilution [11,13,36]. Available conversion methods, such as infusion rate tables or nomograms [37], designed to simplify infusion rate calculation still remain difficult to use and subject to medication errors. Even small errors either in drug calculation or infusion pump flow rate may have a large detrimental impact on the amount of drug delivered [38-40]. This can be harmful for critical patients and even prove fatal [17]. In a study reporting 41% medication errors in pediatric CPRs, the most common error was drug dosage (65% of cases) [19]. Similarly, we found that using the conventional preparation method resulted in dosing errors in 70% of the preparations. Among them, 30% deviated from the original prescribed dose from 100% up to 5233% (ie, 2-fold to 53-fold overdose errors). Disruptive anxiety and exogenous conditions encountered during resuscitation increase the nurse’s cognitive workload and the risk of errors. The cognitive workload has also been shown to be higher when the task is less familiar, as typically seen during CPR [41], which remains uncommon in PED. For instance, at our PED, level one triage status patients represent 0.5% of the ED visits, a proportion similar to the 0.7% (+/- 0.2) seen in the US PED [1]. The lack of practice due to the rareness of certain pathologies or inexperience among the staff regarding vasoactive drugs preparation may add to the complexity of this phenomenon. Some authors have advocated replacing tasks requiring cognitive load during CPR by automated actions [23,41] as much as possible. By guaranteeing an automated, fast, and reliable conversion and preparation, medication errors were totally avoided using PedAMINES. It should be noted that we observed fewer medication errors in the control arm during the second study period, with only a single critical error of more than 20%. This was may be due to the higher number of vials required to prepare norepinephrine (3 vials) compared with that for dopamine (1 vial) or that nurses were a bit more trained (having just done the first scenario). While they had no impact on TDP and TDD, these additional steps of preparation have perhaps limited the rate of errors by strengthening the controls at each preparation step. As tasks were uniform, the complexity of dopamine and norepinephrine preparation by using either PedAMINES or Frank Shann method was similar and not expected to be responsible for the different rates of errors.

Limitations

This study has some limitations. First, it was conducted during a resuscitation simulation-based scenario. This choice was related to the ethical and organizational difficulties of conducting studies with patients in critical situations. However, several studies have demonstrated the benefit of simulation as an investigative research methodology to answer research questions that otherwise could not be answered during CPR [42]. Interindividual diversity among patients and their diseases make CPR studies hard to standardize. Simulation-based CPR scenarios may overcome these limitations by providing a standardized and controlled environment, detailed feedback analysis of the resuscitation stages using audiovisual recordings, and reproducibility. Therefore, high-fidelity simulation has become essential to study resuscitations skills or technologies. We acknowledge that a simulation cannot reveal whether the intervention improves clinical outcome. However, till date, none of the results obtained from simulation-based CPR studies disagreed with those obtained from studies in real life, confirming our study design choice. Although the survival rate has complex and numerous components, it would be interesting to determine in further studies whether saving time and decreasing medication errors owing to the use of PedAMINES would translate into increase in patient survival in real life.

Second, the lack of immersive realism provided during simulated scenarios might directly affect the assessed outcomes. In this study, we used a high-fidelity manikin simulator, which is currently and widely recognized as providing the most realistic and high-yield immersive environment achievable in emergency training [43]. Realism in our study was reflected by the stress levels experienced by the participants. They quoted the simulation as highly stressful when compared with real CPR situations.

Third, accuracy of intervention delivery times could also limit some CPR studies when assessing “times to” outcomes. Conventional paper-based documentation practices used during resuscitation—simulated or not—are often unreliable and inaccurate, either by misreporting intervention delivery times or by missing the delivery completely [44]. In this study, all drug preparation times were chronologically saved in historic log files and videorecorded, rendering our outcomes highly accurate.

Fourth, many different methods are used in PEDs to prepare and administer infusion drugs such as the Frank Shann method or Broselow pediatric resuscitation medication or infusion guide. They all commonly require the user to perform some kind of calculation that may lead to errors and a longer TDP. As described by many authors, the Broselow pediatric emergency tape and affiliate pediatric resuscitation medication or infusion guide have been used with mixed results in many countries and are somehow inaccurate to predict actual weight in almost 20-30% of children, especially for under- and overweight children [45-52]. This may lead to medical errors due to incorrect dosing selection. In 2007, the American Academy of Pediatrics committee on Pediatric Emergency Medicine acknowledged, “although helpful, the Broselow tape is not ideal” [53]. Frank Shann method offers the possibility to precisely adjust the drug doses on patient's actual weights and is used in many hospitals worldwide. This was the rationale to use it as a comparator in our study.

Finally, our study was not intended to compare PedAMINES with "smart" IV pumps. Standardizing drug concentrations of premixed drug and varying infusion rates with smart pumps implicate to deal with poor dose and rate precision in already unstable and critically ill children. As recently reported, no conclusive evidence shows that smart pumps do prevent medication errors and adverse drug events [54,55]. In addition, little is known about the kind of errors that still occur with their use. Moreover, the number of premixed drugs required, their chemical stability over time [56-59] and the lack of specialized pharmacy facility in many smaller hospitals or in other countries around the world, limit their use. Further studies comparing PedAMINES and smart pumps would be valuable.

Conclusions

In summary, compared with a conventional and internationally used preparation method, we found that a mobile app developed following an evidence-based ergonomic-driven approach dramatically reduced TDP and TDD, as well as the medication error rate. The interindividual variance was also reduced by using the app, suggesting a worthwhile benefit to its use by nurses with different experience levels. A large multicenter randomized trial is further needed to assess this assumption in primary and secondary care hospitals.

Acknowledgments

We thank all the study participants for their contribution. All authors declare no source of funding, no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

Conflicts of Interest

Some authors are the owners of the app PedAMINES that will be available in the near future on the Google Play Store and the Apple Store. The authors therefore declare a direct financial interest to market this app.

Authors' Contributions

JNS and SM did the literature search and reading. JNS wrote the manuscript and prepared the figures and tables. JNS, FE, CC, KH, and SM collected data and carried the statistical analysis. JNS, FE, CL, and SM carried out the development of the project software. FE, CC, LL, KH, OS, AG, CL, and SM were involved in the concept and design of the study and critical review of manuscript content. All authors have contributed to, seen and approved the final, submitted version of the manuscript and had full access to all the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author confirms that he had full access to the participants' data and endorsed the final responsibility for the submission. He further affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. Consent to publication was obtained from the participants.

Multimedia Appendix 1

Frank Shann Drug Doses.

[PNG File, 273KB - jmir_v19i2e31_app1.png]

Multimedia Appendix 2

The table details the time expressed in seconds to drug preparation (TDP) and delivery (TDD) for each nurse and for each drug, by study period.

[PNG File, 2MB - jmir_v19i2e31_app2.png]

Multimedia Appendix 3

The table details the deviation expressed in percent in drug dosage (overdose) from the prescribed dose for each nurse and for each drug, by study period.

[PNG File, 1MB - jmir_v19i2e31_app3.png]

Multimedia Appendix 4

CONSORT-EHEALTH checklist V1.6.1 [60].

[PDF File (Adobe PDF File), 738KB - jmir_v19i2e31_app4.pdf]

References

1. CDC. 2011. National Hospital Ambulatory Medical Care Survey: 2011 Emergency Department Summary Tables URL: https://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2011_ed_web_tables.pdf [accessed 2017-01-22] [WebCite Cache ID 6nh4kQv4M]
2. Gerein RB, Osmond MH, Stiell IG, Nesbitt LP, Burns S, OPALS Study Group. What are the etiology and epidemiology of out-of-hospital pediatric cardiopulmonary arrest in Ontario, Canada? Acad Emerg Med 2006 Jun;13(6):653-658 [FREE Full text] [doi: 10.1197/j.aem.2005.12.025] [Medline: 16670256]
3. Kuisma M, Suominen P, Korpela R. Paediatric out-of-hospital cardiac arrests--epidemiology and outcome. Resuscitation 1995 Oct;30(2):141-150. [Medline: 8560103]
4. Sirbaugh PE, Pepe PE, Shook JE, Kimball KT, Goldman MJ, Ward MA, et al. A prospective, population-based study of the demographics, epidemiology, management, and outcome of out-of-hospital pediatric cardiopulmonary arrest. Ann Emerg Med 1999 Feb;33(2):174-184. [Medline: 9922413]
5. Atkins DL, Everson-Stewart S, Sears GK, Daya M, Osmond MH, Warden CR, Resuscitation Outcomes Consortium Investigators. Epidemiology and outcomes from out-of-hospital cardiac arrest in children: the Resuscitation Outcomes Consortium Epistry-Cardiac Arrest. Circulation 2009 Mar 24;119(11):1484-1491 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.108.802678] [Medline: 19273724]
6. Knudson JD, Neish SR, Cabrera AG, Lowry AW, Shamszad P, Morales DL, et al. Prevalence and outcomes of pediatric in-hospital cardiopulmonary resuscitation in the United States: an analysis of the Kids' Inpatient Database*. Crit Care Med 2012 Nov;40(11):2940-2944. [doi: 10.1097/CCM.0b013e31825feb3f] [Medline: 22932398]
7. Topjian AA, Berg RA, Nadkarni VM. Pediatric cardiopulmonary resuscitation: advances in science, techniques, and outcomes. Pediatrics 2008 Nov;122(5):1086-1098 [FREE Full text] [doi: 10.1542/peds.2007-3313] [Medline: 18977991]
8. Matos RI, Watson RS, Nadkarni VM, Huang H, Berg RA, Meaney PA, American Heart Association’s Get With The Guidelines–Resuscitation (Formerly the National Registry of Cardiopulmonary Resuscitation) Investigators. Duration of cardiopulmonary resuscitation and illness category impact survival and neurologic outcomes for in-hospital pediatric cardiac arrests. Circulation 2013 Jan 29;127(4):442-451. [doi: 10.1161/CIRCULATIONAHA.112.125625] [Medline: 23339874]
9. Andersen LW, Berg KM, Saindon BZ, Massaro JM, Raymond TT, Berg RA, American Heart Association Get With the Guidelines–Resuscitation Investigators. Time to Epinephrine and Survival After Pediatric In-Hospital Cardiac Arrest. J Am Med Assoc 2015 Aug 25;314(8):802-810. [doi: 10.1001/jama.2015.9678] [Medline: 26305650]
10. Hubble MW, Johnson C, Blackwelder J, Collopy K, Houston S, Martin M, et al. Probability of return of spontaneous circulation as a function of timing of vasopressor administration in out-of-hospital cardiac arrest. Prehosp Emerg Care 2015;19(4):457-463. [doi: 10.3109/10903127.2015.1005262] [Medline: 25909945]
11. Adapa RM, Mani V, Murray LJ, Degnan BA, Ercole A, Cadman B, et al. Errors during the preparation of drug infusions: a randomized controlled trial. Br J Anaesth 2012 Nov;109(5):729-734 [FREE Full text] [doi: 10.1093/bja/aes257] [Medline: 22850220]
12. Polischuk E, Vetterly CG, Crowley KL, Thompson A, Goff J, Nguyen-Ha P, et al. Implementation of a standardized process for ordering and dispensing of high-alert emergency medication infusions. J Pediatr Pharmacol Ther 2012 Apr;17(2):166-172 [FREE Full text] [doi: 10.5863/1551-6776-17.2.166] [Medline: 23118669]
13. Larsen GY, Parker HB, Cash J, O'Connell M, Grant MC. Standard drug concentrations and smart-pump technology reduce continuous-medication-infusion errors in pediatric patients. Pediatrics 2005 Jul;116(1):e21-e25. [doi: 10.1542/peds.2004-2452] [Medline: 15995017]
14. Moyen E, Camiré E, Stelfox HT. Clinical review: medication errors in critical care. Crit Care 2008;12(2):208 [FREE Full text] [doi: 10.1186/cc6813] [Medline: 18373883]
15. Kaushal R, Bates DW, Landrigan C, McKenna KJ, Clapp MD, Federico F, et al. Medication errors and adverse drug events in pediatric inpatients. J Am Med Assoc 2001 Apr 25;285(16):2114-2120. [Medline: 11311101]
16. Hoyle JD, Davis AT, Putman KK, Trytko JA, Fales WD. Medication dosing errors in pediatric patients treated by emergency medical services. Prehosp Emerg Care 2012;16(1):59-66. [doi: 10.3109/10903127.2011.614043] [Medline: 21999707]
17. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Dtsch Arztebl Int 2012 Sep;109(38):609-616 [FREE Full text] [doi: 10.3238/arztebl.2012.0609] [Medline: 23093991]
18. Gonzales K. Medication administration errors and the pediatric population: a systematic search of the literature. J Pediatr Nurs 2010 Dec;25(6):555-565. [doi: 10.1016/j.pedn.2010.04.002] [Medline: 21035020]
19. Porter E, Barcega B, Kim TY. Analysis of medication errors in simulated pediatric resuscitation by residents. West J Emerg Med 2014 Jul;15(4):486-490 [FREE Full text] [doi: 10.5811/westjem.2014.2.17922] [Medline: 25035756]
20. Ehrler F, Haller G, Sarrey E, Walesa M, Wipfli R, Lovis C. Assessing the usability of six data entry mobile interfaces for caregivers: a randomized trial. JMIR Hum Factors 2015 Dec 15;2(2):e15 [FREE Full text] [doi: 10.2196/humanfactors.4093] [Medline: 27025648]
21. Hagberg H, Siebert J, Gervaix A, Daehne P, Lovis C, Manzano S, et al. Improving drugs administration safety in pediatric resuscitation using mobile technology. Stud Health Technol Inform 2016;225:656-657. [Medline: 27332296]
22. Shann F. Drug Doses RCH Intensive Care Unit, 16th Edition. Melbourne: Parkville, Victoria Collective Pty. Ltd; 2014.

23. Moreira ME, Hernandez C, Stevens AD, Jones S, Sande M, Blumen JR, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated emergency department pediatric resuscitations. Ann Emerg Med 2015 Aug;66(2):97-106.e3 [FREE Full text] [doi: 10.1016/j.annemergmed.2014.12.035] [Medline: 25701295]
24. sealedenvelope. Randomisation and Online Databases for Clinical Trials URL: https://www.sealedenvelope.com [accessed 2016-11-17] [WebCite Cache ID 6m4ZYjUL5]
25. Lutz N, Vandermensbrugghe NG, Dolci M, Amiet V, Racine L, Carron P. Pediatric emergencies admitted in the resuscitation room of a Swiss university hospital. Pediatr Emerg Care 2014 Oct;30(10):699-704. [doi: 10.1097/PEC.0000000000000231] [Medline: 25272075]
26. Nolan JP, Neumar RW, Adrie C, Aibiki M, Berg RA, Bbttiger BW, International Liaison Committee on Resuscitation, Emergency Cardiovascular Care Committee, American Heart Association, Council on Cardiovascular Surgery and Anesthesia, Council on Cardiopulmonary, Perioperative, and Critical Care, Council on Clinical Cardiology, Council on Stroke. Post-cardiac arrest syndrome: epidemiology, pathophysiology, treatment, and prognostication: a scientific statement from the International Liaison Committee on Resuscitation; the American Heart Association Emergency Cardiovascular Care Committee; the Council on Cardiovascular Surgery and Anesthesia; the Council on Cardiopulmonary, Perioperative, and Critical Care; the Council on Clinical Cardiology; the Council on Stroke (Part II). Int Emerg Nurs 2010 Jan;18(1):8-28. [doi: 10.1016/j.ienj.2009.07.001] [Medline: 20129438]
27. Checchia PA, Sehra R, Moynihan J, Daher N, Tang W, Weil MH. Myocardial injury in children following resuscitation after cardiac arrest. Resuscitation 2003 May;57(2):131-137. [Medline: 12745180]
28. Donoghue AJ, Nadkarni V, Berg RA, Osmond MH, Wells G, Nesbitt L, CanAm Pediatric Cardiac Arrest Investigators. Out-of-hospital pediatric cardiac arrest: an epidemiologic review and assessment of current knowledge. Ann Emerg Med 2005 Dec;46(6):512-522. [doi: 10.1016/j.annemergmed.2005.05.028] [Medline: 16308066]
29. Nadkarni VM, Larkin GL, Peberdy MA, Carey SM, Kaye W, Mancini ME, National Registry of Cardiopulmonary Resuscitation Investigators. First documented rhythm and clinical outcome from in-hospital cardiac arrest among children and adults. J Am Med Assoc 2006 Jan 4;295(1):50-57. [doi: 10.1001/jama.295.1.50] [Medline: 16391216]
30. Neumar RW, Nolan JP, Adrie C, Aibiki M, Berg RA, Böttiger BW, et al. Post-cardiac arrest syndrome: epidemiology, pathophysiology, treatment, and prognostication. A consensus statement from the International Liaison Committee on Resuscitation (American Heart Association, Australian and New Zealand Council on Resuscitation, European Resuscitation Council, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Council of Asia, and the Resuscitation Council of Southern Africa); the American Heart Association Emergency Cardiovascular Care Committee; the Council on Cardiovascular Surgery and Anesthesia; the Council on Cardiopulmonary, Perioperative, and Critical Care; the Council on Clinical Cardiology; and the Stroke Council. Circulation 2008 Dec 2;118(23):2452-2483 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.108.190652] [Medline: 18948368]
31. López-Herce J, del Castillo J, Matamoros M, Canadas S, Rodriguez-Calvo A, Cecchetti C, Iberoamerican Pediatric Cardiac Arrest Study Network RIBEPCI. Post return of spontaneous circulation factors associated with mortality in pediatric in-hospital cardiac arrest: a prospective multicenter multinational observational study. Crit Care 2014 Nov 03;18(6):607 [FREE Full text] [doi: 10.1186/s13054-014-0607-9] [Medline: 25672247]
32. Topjian AA, French B, Sutton RM, Conlon T, Nadkarni VM, Moler FW, et al. Early postresuscitation hypotension is associated with increased mortality following pediatric cardiac arrest. Crit Care Med 2014 Jun;42(6):1518-1523 [FREE Full text] [doi: 10.1097/CCM.0000000000000216] [Medline: 24561563]
33. de Caen AR, Berg MD, Chameides L, Gooden CK, Hickey RW, Scott HF, et al. Part 12: pediatric advanced life support: 2015 American heart association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2015 Nov 3;132(18 Suppl 2):S526-S542 [FREE Full text] [doi: 10.1161/CIR.0000000000000266] [Medline: 26473000]
34. Rinke ML, Bundy DG, Velasquez CA, Rao S, Zerhouni Y, Lobner K, et al. Interventions to reduce pediatric medication errors: a systematic review. Pediatrics 2014 Aug;134(2):338-360 [FREE Full text] [doi: 10.1542/peds.2013-3531] [Medline: 25022737]
35. Kozer E, Seto W, Verjee Z, Parshuram C, Khattak S, Koren G, et al. Prospective observational study on the incidence of medication errors during simulated resuscitation in a paediatric emergency department. Br Med J 2004 Dec 4;329(7478):1321 [FREE Full text] [doi: 10.1136/bmj.38244.607083.55] [Medline: 15454495]
36. Ameer A, Dhillon S, Peters M, Ghaleb M. Systematic literature review of hospital medication administration errors in children. Integr Pharm Res Pract 2015 Nov:153. [doi: 10.2147/IPRP.S54998]
37. Sherry E, Burton GW, Wilkins DG. Infusion nomograms. Anaesthesia 1993 May;48(5):396-401 [FREE Full text] [Medline: 8317648]
38. Allen EM, Van Boerum DH, Olsen AF, Dean JM. Difference between the measured and ordered dose of catecholamine infusions. Ann Pharmacother 1995 Nov;29(11):1095-1100. [Medline: 8573951]
39. Lehmann CU, Kim GR, Gujral R, Veltri MA, Clark JS, Miller MR. Decreasing errors in pediatric continuous intravenous infusions. Pediatr Crit Care Med 2006 May;7(3):225-230. [doi: 10.1097/01.PCC.0000216415.12120.FF] [Medline: 16575355]

40. Parshuram CS, To T, Seto W, Trope A, Koren G, Laupacis A. Systematic evaluation of errors occurring during the preparation of intravenous medication. CMAJ 2008 Jan 1;178(1):42-48 [FREE Full text] [doi: 10.1503/cmaj.061743] [Medline: 18166730]
41. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med 2002 Aug;9(8):840-847 [FREE Full text] [Medline: 12153892]
42. Cheng A, Auerbach M, Hunt EA, Chang TP, Pusic M, Nadkarni V, et al. Designing and conducting simulation-based research. Pediatrics 2014 Jun;133(6):1091-1101 [FREE Full text] [doi: 10.1542/peds.2013-3267] [Medline: 24819576]
43. Sahu S, Lata I. Simulation in resuscitation teaching and training, an evidence based practice review. J Emerg Trauma Shock 2010 Oct;3(4):378-384 [FREE Full text] [doi: 10.4103/0974-2700.70758] [Medline: 21063561]
44. Peace JM, Yuen TC, Borak MH, Edelson DP. Tablet-based cardiac arrest documentation: a pilot study. Resuscitation 2014 Feb;85(2):266-269 [FREE Full text] [doi: 10.1016/j.resuscitation.2013.10.013] [Medline: 24157630]
45. Chavez H, Peterson RE, Lo K, Arel M. Weight estimation in an inner-city pediatric ED: the effect of obesity. Am J Emerg Med 2015 Oct;33(10):1364-1367. [doi: 10.1016/j.ajem.2015.07.040] [Medline: 26321165]
46. Clark M, Lewis R, Fleischman R, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, “The hungriest place on earth”. Acad Emerg Med 2016 Jan;23(1):21-28 [FREE Full text] [doi: 10.1111/acem.12854] [Medline: 26671318]
47. Hofer C, Ganter M, Tucci M, Klaghofer R, Zollinger A. How reliable is length-based determination of body weight and tracheal tube size in the paediatric age group? The Broselow tape reconsidered. Br J Anaesth 2002 Feb;88(2):283-285 [FREE Full text] [Medline: 11878662]
48. Knight J, Nazim M, Riggs D, Channel J, Mullet C, Vaughan R, et al. Is the Broselow tape a reliable indicator for use in all pediatric trauma patients? A look at a rural trauma center. Pediatr Emerg Care 2011 Jun;27(6):479-482. [doi: 10.1097/PEC.0b013e31821d8559] [Medline: 21629149]
49. Milne WK, Yasin A, Knight J, Noel D, Lubell R, Filler G. Ontario children have outgrown the Broselow tape. CJEM 2012 Jan;14(1):25-30. [Medline: 22417954]
50. Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF. Use of the Broselow tape may result in the underresuscitation of children. Acad Emerg Med 2006 Oct;13(10):1011-1019 [FREE Full text] [doi: 10.1197/j.aem.2006.06.042] [Medline: 17015417]
51. Sinha M, Lezine MW, Frechette A, Foster KN. Weighing the pediatric patient during trauma resuscitation and its concordance with estimated weight using Broselow Luten Emergency Tape. Pediatr Emerg Care 2012 Jun;28(6):544-547. [doi: 10.1097/PEC.0b013e318258ac2e] [Medline: 22653453]
52. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013 Feb;84(2):227-232. [doi: 10.1016/j.resuscitation.2012.05.028] [Medline: 22796544]
53. Committee on Pediatric Emergency Medicine, American Academy of Pediatrics, Krug S, Frush K. Patient safety in the pediatric emergency care setting. Pediatrics 2007 Dec;120(6):1367-1375. [doi: 10.1542/peds.2007-2902] [Medline: 18055687]
54. Franklin BD. 'Smart' intravenous pumps: how smart are they? BMJ Qual Saf 2017 Feb;26(2):93-94. [doi: 10.1136/bmjqs-2016-005302] [Medline: 27055758]
55. Schnock KO, Dykes PC, Albert J, Ariosto D, Call R, Cameron C, et al. The frequency of intravenous medication administration errors related to smart infusion pumps: a multihospital observational study. BMJ Qual Saf 2017 Feb;26(2):131-140. [doi: 10.1136/bmjqs-2015-004465] [Medline: 26908900]
56. Carr RR, Decarie D, Ensom MH. Stability of epinephrine at standard concentrations. Can J Hosp Pharm 2014 May;67(3):197-202 [FREE Full text] [Medline: 24970939]
57. Dix J, Weber RJ, Frye RF, Nolin TD, Mrvos R, Krenzelok E. Stability of atropine sulfate prepared for mass chemical terrorism. J Toxicol Clin Toxicol 2003;41(6):771-775. [Medline: 14677786]
58. Jäppinen A, Turpeinen M, Kokki H, Rasi A, Ojanen T, Pelkonen O, et al. Stability of sufentanil and levobupivacaine solutions and a mixture in a 0.9% sodium chloride infusion stored in polypropylene syringes. Eur J Pharm Sci 2003 May;19(1):31-36. [Medline: 12729859]
59. Walker SE, Law S, Garland J, Fung E, Iazzetta J. Stability of norepinephrine solutions in normal saline and 5% dextrose in water. Can J Hosp Pharm 2010 Mar;63(2):113-118 [FREE Full text] [Medline: 22478966]
60. 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]

Abbreviations

CPR: cardiopulmonary resuscitation
ED: emergency department
PED: pediatric emergency department

PedAMINES: Pediatric Accurate Medication in Emergency Situations
TDD: time to drug delivery
TDP: time to drug preparation

Edited by G Eysenbach; submitted 17.11.16; peer-reviewed by H Baloch, K Singh, A Landman; comments to author 28.12.16; accepted 11.01.17; published 01.02.17.

Please cite as:
Siebert JN, Ehrler F, Combescure C, Lacroix L, Haddad K, Sanchez O, Gervaix A, Lovis C, Manzano S
A Mobile Device App to Reduce Time to Drug Delivery and Medication Errors During Simulated Pediatric Cardiopulmonary Resuscitation: A Randomized Controlled Trial
J Med Internet Res 2017;19(2):e31
URL: http://www.jmir.org/2017/2/e31/
doi:10.2196/jmir.7005
PMID:28148473

Original Paper

Smartphone-Based Self-Assessment of Stress in Healthy Adult Individuals: A Systematic Review

Helga Þórarinsdóttir[1], B Med; Lars Vedel Kessing[1], MD, Prof. Dr. med.; Maria Faurholt-Jepsen[1], MD

[1] Psychiatric Center Copenhagen, Rigshospitalet, Department O, Copenhagen, Denmark

Corresponding Author:
Maria Faurholt-Jepsen, MD
Psychiatric Center Copenhagen, Rigshospitalet, Department O
Blegdamsvej 9
Copenhagen, 2100
Denmark
Phone: 45 38647073
Fax: 45 386467073
Email: maria@faurholt-jepsen.dk

Abstract

Background: Stress is a common experience in today’s society. Smartphone ownership is widespread, and smartphones can be used to monitor health and well-being. Smartphone-based self-assessment of stress can be done in naturalistic settings and may potentially reflect real-time stress level.

Objective: The objectives of this systematic review were to evaluate (1) the use of smartphones to measure self-assessed stress in healthy adult individuals, (2) the validity of smartphone-based self-assessed stress compared with validated stress scales, and (3) the association between smartphone-based self-assessed stress and smartphone generated objective data.

Methods: A systematic review of the scientific literature was reported and conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement. The scientific databases PubMed, PsycINFO, Embase, IEEE, and ACM were searched and supplemented by a hand search of reference lists. The databases were searched for original studies involving healthy individuals older than 18 years, measuring self-assessed stress using smartphones.

Results: A total of 35 published articles comprising 1464 individuals were included for review. According to the objectives, (1) study designs were heterogeneous, and smartphone-based self-assessed stress was measured using various methods (e.g., dichotomized questions on stress, yes or no; Likert scales on stress; and questionnaires); (2) the validity of smartphone-based self-assessed stress compared with validated stress scales was investigated in 3 studies, and of these, only 1 study found a moderate statistically significant positive correlation ($r=.4$; $P<.05$); and (3) in exploratory analyses, smartphone-based self-assessed stress was found to correlate with some of the reported smartphone generated objective data, including voice features and data on activity and phone usage.

Conclusions: Smartphones are being used to measure self-assessed stress in different contexts. The evidence of the validity of smartphone-based self-assessed stress is limited and should be investigated further. Smartphone generated objective data can potentially be used to monitor, predict, and reduce stress levels.

(J Med Internet Res 2017;19(2):e41) doi:10.2196/jmir.6397

KEYWORDS

smartphone; emotional stress; healthy individuals; self-report; objective smartphone generated measures of stress

Introduction

Many people experience stress, in one form or another, throughout their lives. Stress can be defined as “a state, which is accompanied by physical, psychological or social complaints or dysfunctions and which results from individuals feeling unable to bridge a gap with the requirements or expectations placed on them” [1]. Overall, stress can be divided into 2 types: acute and chronic. Acute stress results from a specific event or situation, is short-lived, and can be accompanied by physical symptoms such as a quickening heartbeat, sweating, and headaches, but can also create motivation to deal with whatever is causing the stress. Chronic stress is the response to prolonged pressure and can stem from traumatic experiences or from the wear and tear of daily stress over a longer time period [2]. Work is the most common cause of stress in the Western world, and more than 1 in every 5 European workers feel stressed [3], whereas 65% of Americans state that they are stressed because of their work [4]. Chronic stress causes overexposure of the body to cortisol and other stress hormones and can be a risk factor for developing diseases. Chronic stress has been associated with cardiovascular problems [5], gastrointestinal problems [6], depression [7], and other psychiatric illnesses [8].

People suffering from chronic stress may be less likely to notice whether they are under high stress at a given time point. Using self-assessment of stress during a time period could potentially increase awareness of stressors and encourage behavioral changes.

In 2015, there were 3.4 billion smartphone subscriptions in the world [9], and it has been estimated that by the year of 2017, one-third of the world's population will use a smartphone [10]. Smartphones can be used for communication, banking, games, looking up information on the Internet, and so forth. During recent years, there has been a growth in the use of smartphones for health monitoring; a search for "health monitor" in Apple's app store alone yields more than 350 results. Smartphone apps can be used to monitor physical activity, calorie intake, sleep quality, the menstrual cycle, and other issues related to health and well-being [11]. Furthermore, monitoring can take place automatically through the sensors embedded within the smartphone, such as accelerometer and microphone, whereas others require that the users interact with the app to register data [12].

Subjective self-assessed stress can be measured using smartphones via ecological momentary assessment (EMA). EMA is a collection of methods used to collect "assessments of subjects' current or recent states, sampled repeatedly over time, in their natural environment" [13]. Advantages of using EMA such as minimization of recall bias [14] and collection of fine-grained real-life data collected during non-laboratory settings have been addressed [15]. Subjective stress can be assessed throughout the day using a time-based EMA where people are prompted to rate or answer questions about their "current stress level" [16]. During recent years, the use of smartphones has been explored within bipolar disorder [17-19], depression [20], and anxiety [21].

Many people carry their smartphones with them throughout the day and are used to interacting with it in many locations, in many situations, and at all times [22]. Thus, smartphone-based data could potentially reflect a person's real-time stress level. Combining smartphone-based self-assessed stress measured by EMA with other smartphone data could help to understand stress better, both on an individual level and on a group level.

However, with no systematic review within this area, the extent to which smartphone-based self-assessed stress has been monitored and evaluated in healthy individuals is unknown. Furthermore, the validity of smartphone-based self-assessed stress compared with other validated stress scales has not been evaluated systematically. Thus, the objectives of this systematic review were to evaluate (1) the use of smartphones to measure self-assessed stress in healthy adult individuals, (2) the validity of smartphone-based self-assessed stress compared with validated stress scales, and (3) the association between smartphone-based self-assessed stress and smartphone generated objective data.

This was the first systematic review of smartphone-based self-assessed stress in healthy adult individuals.

Methods

Overview

This systematic review was conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [23]. Methods of the review process and eligibility criteria were established in advance and documented in a review protocol that can be retrieved from the authors upon request. No changes were made to the protocol during the review process.

Eligibility Criteria

Original studies involving healthy individuals older than 18 years measuring self-assessed stress on a smartphone were eligible for review. The language of publication was restricted to English. Papers not meeting eligibility criteria or only describing the technical part of the self-assessment of stress were excluded from review. Where multiple articles were reported on the same study, the article presenting the largest and most detailed dataset was included for review. Only studies in which self-assessed stress was reported on smartphones were eligible for review.

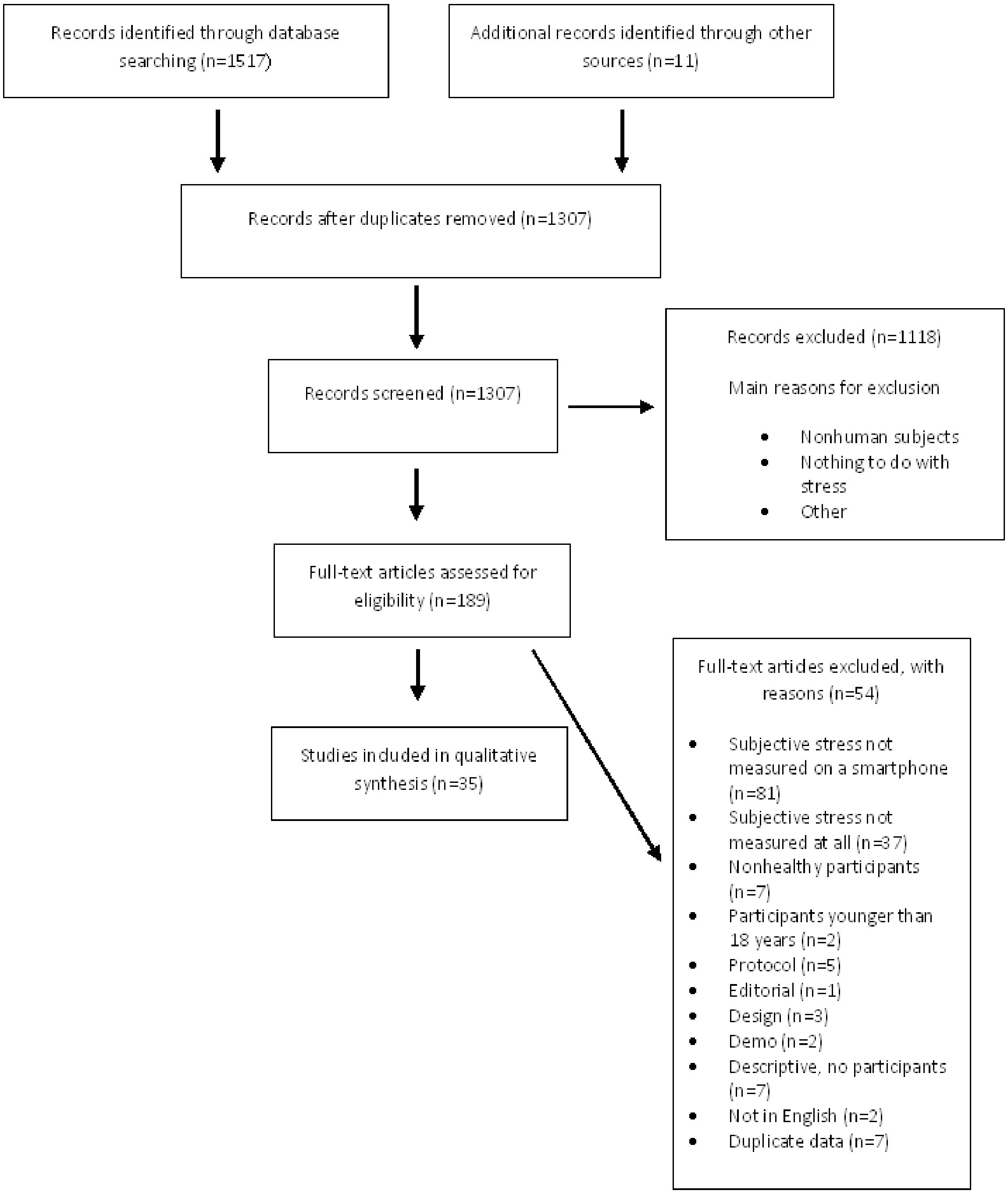
Information Sources and Search Strategy

Published studies were identified by conducting a systematic literature search through the electronic databases PubMed, PsycINFO, Embase, IEEE, and ACM. The literature search was supplemented by a hand search of reference lists of retrieved articles. The literature search was conducted by 1 researcher (HP), without time restrictions, using the following keywords: (stress or psychological stress or emotional stress) AND (smartphone or cell phone or cellular phone or mobile phone or mobile application or ecological momentary assessment or experience sampling method) and covered a period from 1980 to May 2016. The last literature search was conducted on May 4, 2016.

Study Selection and Data Extraction

A PRISMA flow diagram of the study selection process is presented in Figure 1. All identified titles and abstracts were screened for eligibility by 1 researcher (HP). Potentially relevant articles were retrieved and full-text articles then checked for fulfilling eligibility independently by 2 researchers (HP and MFJ). One researcher extracted data (HP), and a second reviewer (MFJ) independently checked the extracted data. Any disagreements were resolved by a discussion between 3 researchers (HP, MFJ, and LVK).

Figure 1. Flow diagram of literature search according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).



Results

Study Selection

The literature search identified a total of 1517 articles from the 5 databases, and 11 additional studies were identified by hand search of reference lists. Removing duplicates left 1307 articles for further evaluation. Reviewing abstracts and titles resulted in the exclusion of a total of 1118 articles for not meeting eligibility criteria, the 2 main reasons for exclusion being not including human subjects and not involving stress. Thus, 189 full-text articles were evaluated for eligibility. Of these, 154 articles were excluded from the review for various reasons (Figure 1), with the main reasons being (1) subjective stress not measured on a smartphone (n=81) and (2) subjective stress not measured at all (n=37). A list of excluded articles can be retrieved from the authors upon request. Thus, a total of 35 articles fulfilled the eligibility criteria and were included for review [24-58].

Study Characteristics

Of the 35 studies, 17 were from the United States and the remainder were from Finland (n=3), Italy (n=4), Germany (n=2), Switzerland (n=2), the United Kingdom (n=3), Australia (n=1), Hong Kong (n=1), Portugal (n=1), and Sweden (n=1). The majority of studies were prospective observational studies [24,28,29,31,33-38,41-45,49-58], 2 were randomized control trials [39,47], 4 were other types of intervention studies [27,40,46,48], 2 were case reports [25,26], and 1 was a cross-sectional study [32]. The study period ranged from 1 hour to 191 days. All the studies were published recently, with the oldest one published in 2007 [44] and more than half of the studies published since 2013 [24,25,27,29-34,36,37,40,42,45,47-54,56-58]. More than half of the studies were published in conference proceedings (n=19), whereas 16 studies were published in peer-reviewed scientific journals (Table 1).

Table 1. Characteristics of studies on smartphone-based self-assessed stress in healthy adult individuals included for systematic review (Studies: N=35).

Author	Publication year	Publication type	Study design	Study location	Study duration (days)	Number of participants, context of assessment	Method for self-assessment of stress	Times per day stress measured	Smartphone operating system
Adams et al [24]	2010	Conference paper	Cohort	United States	10	7, daily life	Taylor 5-item measure	Multiple	Android
Atz [25]	2013	Journal article	Case report	United Kingdom	56	1, daily life	7-point Likert scale	Multiple	iOS
Ayzenberg et al [26]	2012	Conference paper	Case report	United States	8.3	1, daily life	7-point Likert scale	N/A	N/A
Bandiera et al [27]	2016	Journal article	Interventional	United States	14	139, smoking cessation	5-point Likert scale	5	Android
Berndt et al [28]	2011	Conference paper	Cohort	Germany	1	50, daily life	0-100 scale	Multiple	N/A
Bogomolov et al [29]	2014	Conference paper	Cohort	United States	190	117, daily life	7-point scale	1	Android
Carroll et al [30]	2013	Conference paper	Cohort	United States	4	12, emotional eating	Russel Circumplex model	Multiple	Windows
Ceja et al [31]	2015	Journal article	Cohort	Italy	40	30, workplace stress	5-point scale	3	Android
Ciman et al [32]	2015	Conference paper	Cross-sectional	Switzerland	0.04 (1 hour)	13, laboratory	5-point Likert scale	N/A	Android
Ferdous et al [33]	2015	Conference paper	Cohort	Italy	42	28, workplace stress	5-point scale	3	Android
Gaggioli et al [34]	2011	Journal article	Cohort	Italy	7	6, daily life	10-point Likert scale	Multiple	Windows
Gomes et al [35]	2012	Conference paper	Cohort	Portugal	191	5, workplace stress	Questionnaire	N/A	Android
Huang et al [36]	2015	Conference paper	Cohort	United States	28	14, daily life	N/A[a]	N/A	Android
Huh et al [37]	2014	Journal article	Cohort	United States	7	26, smoking behavior	Perceived stress	5	Android
Jin et al [38]	2012	Conference paper	Cohort	Hong Kong	2	30, workplace stress	N/A	Multiple	Android
Kennedy et al [39]	2011	Journal article	Interventional	United Kingdom	33	198, vitamin intake	VAS[b]	2	Other
Lachmann et al [40]	2016	Journal article	Interventional	Sweden	14	33, interprofessional learning	7-point Likert scale	5	N/A
Madan et al [41]	2010	Conference paper	Cohort	United States	73	70, epidemiology	Yes or no	1	Windows
Muaremi et al [42]	2013	Journal article	Cohort	Switzerland	112	35, workplace stress	Continuous response value	5	iOS
Muukkonen et al [43]	2008	Conference paper	Cohort	Finland	14	55, studying	Yes or no	5	Symbian
Muukkonen et al [44]	2007	Conference paper	Cohort	Finland	14	8, studying	Yes or no	5	Symbian
Ottaviani et al [45]	2015	Journal article	Cohort	Italy	1	42, daily life	Yes or no	Multiple	Android
Parkka et al [46]	2009	Journal article	Interventional	Finland	56	17, workplace stress	Sliding scale	1	Symbian
Pipingas et al [47]	2013	Journal article	Interventional	Australia	112	38, vitamin intake	VAS	0.14 (once a week)	N/A

Author	Publication year	Publication type	Study design	Study location	Study duration (days)	Number of participants, context of assessment	Method for self-assessment of stress	Times per day stress measured	Smartphone operating system
Reitzel et al [48]	2014	Journal article	Interventional	United States	13	22, smoking cessation	5-point Likert scale	5	Android
Sano et al [49]	2015	Conference paper	Cohort	United States	30	66, daily life	Calmness	2	Android
Sano and Picard [50]	2013	Conference paper	Cohort	United States	5	18, daily life	0-100 scale	2	Android
Sarker et al [51]	2014	Conference paper	Cohort	United States	7	30, daily life	6-point scale	Multiple	N/A
Vhaduri et al [52]	2014	Journal article	Cohort	United States	7	30, driving	6-point Likert scale	Multiple	Android
Wang et al [53]	2014	Conference paper	Cohort	United States	70	48, daily life	Taylor 5-item measure	Multiple	Android
Weppner et al [54]	2013	Conference paper	Cohort	Germany	84	9, daily life	10-point Likert scale	10	Android
Witiewitz et al [55]	2012	Journal article	Cohort	United States	21	86, concurrent drinking and smoking	5-point Likert scale	3	N/A
Wray et al [56]	2015	Journal article	Cohort	United States	14	76, smoking behavior	5-point Likert scale	4	iOS
Zenk et al [57]	2014	Journal article	Cohort	United States	7	100, snack-food intake	Yes or no	5	Android
Zenonos et al [58]	2016	Conference paper	Cohort	United Kingdom	11	4, workplace stress	0-100 scale	Multiple	Android

[a] N/A: not available.

[b] VAS: visual analog scale.

Study Participants

Overall, the studies comprised a total of 1464 healthy adult participants, with sample sizes in individual studies varying from 1 to 198 participants. The mean age of the participants was available for 19 of the studies and ranged from 20.1-52.47 years [27,31-34,37-39,43-46,48-52,55,57]. Gender distribution was available for 25 studies [24,27,30-34,36,37,39-42,45-53,55-57], and of these, 4 studies had equal gender distribution [34,47,51,52], whereas there were 2 large gender-specific studies, 1 male [39] and 1 female [57]. In 9 of the studies, the participants were exclusively students [36,40,41,43,44,49,51-53,55], and in 4 studies, the participants were exclusively employees [31,33,42,58].

Smartphones

The majority of the studies used Android-based smartphones (n=19), 3 used Windows-based smartphones [30,34,41], 3 studies used iPhones [25,49,56], 4 studies [39,43,44,46] used other types of smartphones, whereas the remaining 6 studies did not specify what type of smartphones or operating systems were used [26,28,40,47,51,55]. In 15 of the studies, smartphones were provided for the participants, whereas participants used their own smartphones in 4 studies [32,41,42,49]. Although some participants used their own smartphone in 4 studies, other participants borrowed a smartphone [24,37,53,56]. In total, 12 studies did not specify ownership of the smartphones used [25,26,28,30,39,43,44,50,51,55,57,58].

Self-Assessed Stress

Overall, the included studies used many different methods to measure smartphone-based self-assessed stress. The most common method (n=11) was using a Likert scale (from a 5-point scale to a 10- or 100-point scale) [25-28,32,34,40,48,50,52,54-56,58]. Five studies used a yes or no answer question to measure self-assessed stress [41,43-45,57], and 5 studies used questionnaires [24,49,50,53]. Two studies did not specify how smartphone-based self-assessed stress was measured [36,38].

The frequency of smartphone-based self-assessed stress reports varied . In most of the studies, participants were asked to report their stress levels multiple times per day: from twice a day [39,49,50] to up to once every half hour [24]. In 1 study, participants reported self-assessed stress on a weekly basis [47], whereas in 3 studies, self-assessed stress was reported once per day [29,41,46]. In 4 studies, the frequency of self-assessment was not specified [26,32,35,36].

Context

Six studies investigated self-assessed stress in the context of the workplace [31,33,35,38,42,58], and 1 study in relation to rehabilitation after work-related stress [46]. Two studies measured self-assessed stress in relation to smoking cessation [27,48], 2 in relation to smoking behavior [37,56], and 1 in relation to concurrent smoking and drinking [55]. Two studies investigated self-assessed stress levels in relation to vitamin intake [39,47], 1 in relation to emotional eating [30], and another in relation to snack-food intake [57]. Three studies looked at self-assessed stress in the context of studying [40,43,44]. One study was done in a laboratory context [32], another in relation to driving [52], and a third looked at stress levels in epidemiological behavior context [41]. The remaining studies (n=13) reported no specific context, and participants registered self-assessed stress during their everyday life. About half (n=16) of the studies investigated stress as the primary objective [24-26,28,29,31-33,35,38,42,50,52,54,58].

Validity of Smartphone-Based Self-Assessed Stress

In 5 studies, validated stress scales in addition to smartphone-based self-assessed stress were reported. Four of these studies used the Perceived Stress Scale (PSS) [24,49,50,53,59], and 1 used Derogatis Stress Profile (DSP) [46,60]. In 2 of the studies, participants filled out the PSS at baseline only [24,50], and in 2 studies, participants filled out the PSS at both baseline and follow-up [49,53]. The study using DSP was an interventional study, and participants filled out the scale 4 times during the study period [46].

Three studies investigated the correlation between smartphone-based self-assessed stress and validated stress scales [24,46,53]. Adams et al reported a statistically nonsignificant correlation (r=.562, P=.11) between smartphone-based self-assessed stress levels and PSS score [24]. Another study by Parkaa et al reported a statistically nonsignificant correlation (ρ=.07, P=.64) between smartphone-based self-assessed stress and DSP score [46]. Finally, a study by Wang et al reported a statistically significant positive moderate correlation between smartphone-based self-assessed stress and PSS score both pre- (r=.458, P=.003) and poststudy (r=.412, P=.009) [53].

Smartphone Generated Objective Data

A total of 13 studies collected smartphone generated objective data [24,26,29,31,33,41,42,48-53]. Six studies investigated the association between smartphone generated objective data and smartphone-based self-assessed stress [24,29,31,33,41,49]. Among these, 2 studies investigated the association between smartphone-based self-assessed stress and verbal data [24,33]; Adams et al reported a statistically positive correlation (r=.59, P value not specified) between smartphone-based self-assessed stress and voice-stress, whereas Ferdous et al reported a significant positive correlation between smartphone-based self-assessed stress and duration of verbal interaction for 17 of their 28 participants (r=.06-.55, P<.005).

A study by Madan et al reported that communication diversity was reduced for participants who often assessed themselves as being stressed, and the authors interpreted this as a tendency to isolate [41]. A study by Sano et al reported that higher self-assessed stress levels were statistically significantly correlated with lower activity level in the evening, fewer and shorter text messages sent, and less screen activity in the evening [49].

Two studies investigated the association between self-assessed stress and smartphone generated objective data in order to detect stress [29,31]. A study by Ceja et al looked at smartphone generated objective data from the accelerometer and “achieved a maximum overall accuracy of 71% for user-specific models and an accuracy of 60% for the use of similar-users models” to classify self-assessed stress levels [31]. A study by Bogomolov et al collected both social features (phone calls and text messages) and proximity features (Bluetooth) and obtained “the accuracy score of 72.28% for a 2-class daily stress recognition problem” [29].

Discussion

Principal Findings

This was the first systematic review on smartphone-based self-assessment of stress in healthy adult individuals. A total of 35 published articles involving a total of 1464 participants were included for review. Overall, the study designs were highly heterogeneous, using various methods of self-assessment in different contexts. Most of the studies were conducted in the United States or Western Europe. Android-based smartphones were most commonly used for measuring self-assessed stress, many participants borrowed smartphones during the studies, and often stress was reported multiple times per day.

Regarding the validity of smartphone-based self-assessed stress levels, stress levels measured using validated stress scales were available in 5 studies, but only 3 of these studies investigated the correlation between smartphone-based self-assessed stress and validated stress scales. Among these 3 studies, only 1 study found a statistically significant positive correlation between self-assessed stress and a validated stress scale (PSS) [53]. It should be noted that the study by Wang et al included a larger sample (n=48) compared with the other 2 studies combined (n=7; n=17) [24,46], suggesting a low statistical power of the other 2 studies. In addition, the study by Wang et al included university students on a university campus, limiting the generalizability of the study findings. The validity of smartphone-based self-assessment of stress may be different across populations and should be investigated further in future studies. Thus, findings from this systematic review suggest that the validity of smartphone-based self-assessed stress has been sparingly investigated and is unknown. The studies included described convergent validity of smartphone-based self-assessment of stress. Other parameters such as sleep, mood, and activity level may correlate with validated stress scales; however, content validity was not investigated in this review. In addition, the reliability and predictive validity of smartphone-based self-assessment of stress were also not investigated.

Smartphone generated objective data were collected in 13 studies and 6 studies investigated the association between smartphone-based self-assessed stress and these objective data. Two studies found a positive correlation between self-assessed stress and verbal data, whereas another 2 studies found associations between self-assessed stress and communication diversity, activity levels, text messages, and screen on or off patterns. The last 2 studies found smartphone generated data to be a predictor (accuracy up to 72.28%) for detecting self-assessed stress. Overall, regarding smartphone generated objective data, the studies collected various smartphone generated data and the results seem exploratory, with a tendency to report statistically significant positive correlations with self-assessed stress only.

A majority of the included studies collected objective data alongside the self-assessed data. Some of them used physiological measures collected from various worn sensors, but others only used objective data collected from sensors embedded within the smartphones. Seven studies collected all 3 kinds of stress measures. Collecting physiological measures such as heart rate requires participants to carry additional sensors (user burden), whereas smartphone generated objective data are collected from a smartphone that is most likely already being carried around. Smartphone generated objective data can usually be collected automatically, eliminating attrition due to monitoring. Objective smartphone data are behavioral data that can reflect behavior related to stress. Different people react differently to stress, and combined with self-assessed data on stress, smartphone generated objective data might be used for detecting stress. Early stress detection in healthy populations such as students and employees could help to prevent stress-related diseases. Thus, the use of smartphone generated objective data as a marker of stress in healthy individuals has been sparingly investigated and future well-designed studies investigating this would be interesting.

Stress levels were assessed from self-reported data, both from smartphones and from validated scales. PSS was developed in 1983. It has 10 questions and is widely used within psychological and psychiatric sciences. It has shown good internal reliability (Cronbach alpha=.78-.91 [61]) and is correlated with various self-report and behavioral criteria [59]. DSP is a 77-item self-report inventory developed in 1980 and has also shown good internal reliability (Cronbach alpha=.83-.88 [62]). It should be emphasized that the different methods for self-assessment of stress, smartphone-based and validated scales, do not necessarily measure the same thing. Validated stress scales measure more long-term stress levels, whereas self-assessment on smartphones is more about current stress levels. Validated scales such as PSS have a somewhat clear definition of stress, as they have several items that the participants have to answer. Many of the smartphone-based self-assessment measures of stress were not explicit in their definition of stress, and participants often only answered 1 question about their level of stress. Stress is a popular term and can mean different things to different people; some people might only register stress that they experience as a negative thing, whereas others might also register the kind of stress (eustress) that is positive and can be motivating. As noted in a study by Muaremi et al, stress was not necessarily a negative event or feeling for some of the participants [42].

Registering self-assessed stress multiple times a day can be a tool to help people self-monitor stress levels. In this way, self-monitoring may play a role in helping people to manage stress. Self-monitoring brings awareness of stress levels and encourages behavioral change according to a situation [63]. However, being asked to self-assess one’s stress level up to multiple times a day could introduce a negativity bias. This could result in participants assessing their stress to be higher than it actually is and even potentially cause more stress per se. It may be that measurements in themselves are stressful, but also the situation to have the self-assessed results of chronic stress constantly at hand and to be unable to cope with a given stressful situation. In this case, people may be constantly reminded that they are unable to cope with stress, which may be the reason they are measuring self-assessed stress in the first place. Investigating the effect of introducing coaching or coping elements to the self-assessment apps would be interesting. It should be stressed that we identified no study that investigated whether the use of smartphone to continuously monitor stress—subjectively reported or objectively assessed—per se had a reducing effect on stress level. Whether self-assessments multiple times a day would be a threat to the reliability and validity is unknown and should be investigated further. Most studies looked at self-assessed stress in everyday life, either without context or in the context of work or studying. Many people carry their smartphones with them during most of the day and therefore smartphones are a device well suited for this type of data collection. Registering stress multiple times a day, in different situations, can shed light on where and when people are experiencing stress.

A study by Wang et al looked at stress in students over a whole semester and revealed how their self-assessed stress level increased as their workload increased, with the peak being during final examinations [53]. Following a group of people prospectively over time could help distinguish between the normal stresses that come and go and the chronic, potentially health-damaging kind of stress. Being aware of chronic stress is the first step toward eradicating or minimizing it.

Most studies measured self-assessed stress on Android-based smartphones, and many participants were provided with smartphones during the study period. Allowing participants to use their own smartphones to collect self-assessment of stress would be the least disruptive for participants, as they are already familiar with the device. Using one’s own smartphone would also be likely to more accurately reflect real life, especially in regards to the objective smartphone data. It is possible that participants did not, in all cases, own smartphones. It is also possible that the study smartphones were specially programmed for the study or that participants’ smartphones were different from the ones that were required for the study.

Smartphones constitute a new and an exciting research tool within psychological well-being and health care. Nevertheless, the majority of the identified studies have been published in proceedings from technological conferences. In general, many of these studies focused primarily on the technical side of the smartphone system, and a number of these did not present data on population characteristics such as age [24,26,28-30,35,36,41,44,53,54,58], gender [26,28,29,35,38,43,44,54,58], or employment status of participants [24-30,32,34,35,38,39,45,48,50,54,56].

Limitations

Limitations at a study level: Several concerns regarding the individual studies and outcomes limited the overall findings of this study. The included studies had highly heterogeneous designs and used various methods to measure smartphone-based self-assessed stress. In addition, in many cases studies did not include clear descriptions of the recruitment process. The studies included were at risk of selection bias, and at an individual study level, there was a lack of information on potential confounding factors such as age, gender, and educational level, which possibly could have affected self-assessed stress level. A large part of the studies included a relatively small sample of participants and reported unadjusted statistical analyses. Validated stress scales were only used in 5 studies out of the 35 studies included. More than half of the included studies did not investigate stress as their primary objective, and information was therefore limited: only 1 out of the 4 largest studies had stress as their primary objective. In general, studies focusing on stress had fewer participants (mean n=24.7) compared with the studies not focusing on stress (mean n=56.3). Self-assessed stress was investigated in selected groups, often recruited through convenience sampling at a university or a workplace. In many of the studies, participants were provided with a smartphone to use during the study period, and some participants received economic incentives to fill out the self-assessments of stress. The generalizability of these studies was therefore limited, but findings could be relevant for more narrow populations such as university students. Overall, methodological limitations related to study designs, self-assessments of stress, as well as statistical analyses of the included studies were observed. There is a need for studies investigating the use and validity of smartphone-based self-assessed stress in more general populations.

Limitations at a review level: Some limitations to this review should be mentioned. Research using smartphones is expanding, and due to the intersectionality of this research (medicine, psychology, and information technology), studies are being published in very diverse forms and places. Our review shows that many of these kinds of studies are being published in conference proceedings. Therefore, conducting a search strategy that is able to capture all relevant scientific articles is a challenge. The review process was restricting among healthy smartphone users and articles published in English, which might have reduced the global acceptance.

Perspectives and Implications

Stress has become a major health problem in the Western world. Awareness of one’s own stress level is important, and smartphones are potentially a proper minimally intrusive tool for self-assessment of stress. Self-assessment of stress using smartphones in everyday life is a step toward stress awareness. Looking at self-reported stress levels in relation to other more objective data from smartphones, such as geolocation and physical activity, could help to further understanding of stress and stress-related behavior. However, well-designed studies using strict methodology investigating the validity of smartphone-based self-assessment of stress are warranted. Future studies should investigate how to validly measure subjective stress using smartphones, which by nature is accurate in time and place, in contrast to a self-reported scale on stress administered once a day or less frequently. They should also collect information on and address possible confounding factors in the statistical analyses. In addition, and of even more paramount importance, they should investigate in a randomized controlled trial whether the use of smartphone to monitor stress—subjectively or objectively assessed—per se has a beneficial or detrimental effect on stress level.

This review included only studies with healthy adult participants. Smartphones can and are also being used to measure self-assessed stress in various patient populations, especially within the mental health field, where stress is a risk factor. However, addressing this aspect was beyond the scope of this review.

Conclusions

This systematic review identified 35 studies using smartphones to measure self-assessed stress in healthy adults. The studies were from different countries and used different self-assessment methods in varying contexts, such as in the workplace, in relation to smoking cessation, and on university campuses. Android-based smartphones were most commonly being used, and the validity of smartphone-based self-assessed stress compared with validated stress scales was limited by low statistical power of the individual studies and small number of studies reporting on validated scales. Some smartphone generated objective data, including voice, activity, and general usage data, were associated with self-assessed stress measured on smartphones. Smartphone generated objective data could represent a potential tool for predicting stress levels. There is a need for further studies investigating the validity of smartphone-based self-assessed stress and smartphone generated objective measures of stress using validated stress scales, and studies investigating the beneficial or detrimental effects of smartphone-based monitoring stress, both subjectively and objectively, on stress levels per se.

Acknowledgments

The study was supported by the Lundbeck Foundation.

Conflicts of Interest

Lars Vedel Kessing has, in the last 3 years, been a consultant for Lundbeck and Astra Zeneca. Maria Faurholt-Jepsen has been a consultant for Eli Lilly and Lundbeck.

References

1. Eurofound.europa. European Foundation for the Improvement of Living and Working Conditions: Work-related stress URL: http://www.eurofound.europa.eu/sites/default/files/ef_files/docs/ewco/tn1004059s/tn1004059s.pdf [WebCite Cache ID 6j5kFS2q2]
2. APA. Stress: The Different kinds of Stress URL: http://www.apa.org/helpcenter/stress-kinds.aspx [accessed 2017-02-06] [WebCite Cache ID 6o4QqX44b]
3. Eurofound.europa. Fourth European Working conditions Survey (2005) URL: http://www.eurofound.europa.eu/ewco/surveys/ewcs2005/index [accessed 2016-07-18] [WebCite Cache ID 6j5ksPc7t]
4. APA. Stress in America: The Impact of Discrimination URL: http://www.apa.org/news/press/releases/stress/2015/impact-of-discrimination.pdf [WebCite Cache ID 6j5lI4YXz]
5. Chockalingam A, Venkatesan S, Dorairajan S, Moorthy C, Chockalingam V, Subramaniam T. Estimation of subjective stress in acute myocardial infarction. J Postgrad Med 2003;49(3):207-210 [FREE Full text] [Medline: 14597781]
6. Mönnikes H, Tebbe JJ, Hildebrandt M, Arck P, Osmanoglou E, Rose M, et al. Role of stress in functional gastrointestinal disorders. Evidence for stress-induced alterations in gastrointestinal motility and sensitivity. Dig Dis 2001;19(3):201-211. [Medline: 11752838]
7. Plieger T, Melchers M, Montag C, Meermann R, Reuter M. Life stress as potential risk factor for depression and burnout. Burn Res 2015 Mar;2(1):19-24. [doi: 10.1016/j.burn.2015.03.001]
8. Dinan TG. Stress: the shared common component in major mental illnesses. Eur Psychiatry 2005 Oct;20(Suppl 3):S326-S328. [Medline: 16459244]
9. Ericsson. Ericsson mobility report URL: https://www.ericsson.com/res/docs/2016/mobility-report/ericsson-mobility-report-feb-2016-interim.pdf [accessed 2016-07-18] [WebCite Cache ID 6j5lRt4Dy]
10. eMarketer. 2014. Smartphone Users Worldwide Will Total 1.75 Billion in 2014 URL: http://www.emarketer.com/Article/Smartphone-Users-Worldwide-Will-Total-175-Billion-2014/1010536 [accessed 2016-07-18] [WebCite Cache ID 6j5lfjd9Y]
11. Nuviun. Health and wellness apps URL: http://nuviun.com/digital-health/health-and-wellness-apps [accessed 2016-07-18] [WebCite Cache ID 6j5mOE5TK]
12. Kailas A, Chong CC, Watanabe F. From mobile phones to personal wellness dashboards. IEEE Pulse 2010;1(1):57-63. [doi: 10.1109/MPUL.2010.937244] [Medline: 20875965]
13. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [Medline: 18509902]
14. Beal DJ, Weiss HM. Methods of ecological momentary assessment in organizational research. Organ Res Methods 2003 Oct;6(4):440-464. [doi: 10.1177/1094428103257361]
15. Scollon CN, Kim-Prieto C, Diener E. Experience sampling: promises and pitfalls, strengths and weaknesses. J Happiness Stud 2003;4(1):5-34. [doi: 10.1023/A:1023605205115]
16. Robbins M, Kubiak T. Ecological Momentary Assessment in Behavioral Medicine. In: The Handbook of Behavioral Medicine. Oxford, UK: John Wiley & Sons, Ltd; Jan 01, 2014.
17. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Margrethe CE, Bardram JE, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. Int J Methods Psychiatr Res 2016 Dec;25(4):309-323. [doi: 10.1002/mpr.1502] [Medline: 27038019]
18. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]
19. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord 2015 Nov;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]
20. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
21. Mayo-Wilson E, Montgomery P. Media-delivered cognitive behavioural therapy and behavioural therapy (self-help) for anxiety disorders in adults. Cochrane Database Syst Rev 2013;9:CD005330. [doi: 10.1002/14651858.CD005330.pub4] [Medline: 24018460]
22. Raento M, Oulasvirta A, Eagle N. Smartphones: an emerging tool for social scientists. Sociol Methods Res 2009 Feb 01;37(3):426-454. [doi: 10.1177/0049124108330005]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
24. Adams P, Rabbi M, Rahman T, Matthews M, Voida A, Gay G, et al. Towards personal stress informatics: Comparing minimally invasive techniques for measuring daily stress in the wild. 2014 Jan 01 Presented at: Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare. . ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). doi>10.4108/icst.pervasivehealth..254959; 2014; ICST. [doi: 10.4108/icst.pervasivehealth.2014.254959]

25. Atz U. Evaluating experience sampling of stress in a single-subject research design. Pers Ubiquit Comput 2012 Apr 17;17(4):639-652. [doi: 10.1007/s00779-012-0512-7]
26. Ayzenberg Y, Rivera JH, Picard R. FEEL: frequent EDAevent logging -- a mobile social interaction stress monitoring system. 2012 Presented at: CHI EA '12 CHI '12 Extended Abstracts on Human Factors in Computing Systems; May 05 - 10, 2012; Austin, Texas, USA. [doi: 10.1145/2212776.2223802]
27. Bandiera FC, Atem F, Ma P, Businelle MS, Kendzor DE. Post-quit stress mediates the relation between social support and smoking cessation among socioeconomically disadvantaged adults. Drug Alcohol Depend 2016 Jun 1;163:71-76 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.03.023] [Medline: 27085499]
28. Berndt R, Takenga M, Kuehn S, Preik P, Stoll N, Thurow K, et al. A scalable and secure Telematics Platform for the hosting of telemedical applications. 2011 Presented at: e-Health Networking Applications and Services (Healthcom), 2011 13th IEEE International Conference; 13-15 June 2011; Columbia, MO, USA p. 121. [doi: 10.1109/HEALTH.2011.6026726]
29. Bogomolov A, Lepri B, Ferron M, Pianesi F, Pentland A. Daily Stress Recognition from Mobile Phone Data, Weather ConditionsIndividual Traits. In: Proceedings of the 22nd ACM international conference on Multimedia. 2014 Presented at: ACM; 2014; Orlando, Florida. [doi: 10.1145/2647868.2654933]
30. Carroll E, Czerwinski M, Roseway A, Kapoor A, Johns P, Rowan K, et al. Food and Mood: Just-in-Time Support for Emotional Eating. 2013 Jan 01 Presented at: Affective Computing and Intelligent Interaction (ACII), Humaine Association Conference; 2013; Geneva, Switzerland. [doi: 10.1109/ACII.2013.48]
31. Garcia-Ceja E, Osmani V, Mayora O. Automatic stress detection in working environments from smartphones' accelerometer data: a first step. IEEE J Biomed Health Inform 2016 Jul;20(4):1053-1060. [doi: 10.1109/JBHI.2015.2446195] [Medline: 26087509]
32. Ciman M, Wac K, Gaggi O. iSensestress: Assessing stress through human-smartphone interaction analysis. 2015 Jan 01 Presented at: Pervasive Computing Technologies for Healthcare (PervasiveHealth), 9th International Conference on; 2015; Barcelona, Spain. [doi: 10.4108/icst.pervasivehealth.2015.259280]
33. Ferdous R, Osmani V, Beltran MJ, Mayora O. Investigating correlation between verbal interactions and perceived stress. Conf Proc IEEE Eng Med Biol Soc 2015 Aug;2015:1612-1615. [doi: 10.1109/EMBC.2015.7318683] [Medline: 26736583]
34. Gaggioli A, Pioggia G, Tartarisco G, Baldus G, Corda D, Cipresso P, et al. A mobile data collection platform for mental health research. Pers Ubiquit Comput 2011 Sep 30;17(2):241-251. [doi: 10.1007/s00779-011-0465-2]
35. Gomes P, Lopes B, Coimbra M. Vital analysis: field validation of a framework for annotating biological signals of first responders in action. Conf Proc IEEE Eng Med Biol Soc 2012;2012:2128-2131. [doi: 10.1109/EMBC.2012.6346381] [Medline: 23366342]
36. Huang Y, Tang Y, Wang Y. Emotion Map: A Location-based Mobile Social System for Improving Emotion AwarenessRegulation. Vancouver, Canada: Association for Computing Machinery, Inc; 2015 Presented at: CSCW 2015 - Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing; Feb 28, 2015; BC, Canada p. 130-142. [doi: 10.1145/2675133.2675173]
37. Huh J, Shin H, Leventhal AM, Spruijt-Metz D, Abramova Z, Cerrada C, et al. Momentary negative moods and being with friends precede cigarette use among Korean American emerging adults. Nicotine Tob Res 2014 Sep;16(9):1248-1254. [doi: 10.1093/ntr/ntu063] [Medline: 24755398]
38. Jin Z, Tang H, Chen D, Zhang Q. deStress: Mobile and remote stress monitoring, alleviation, and management platform. 2012 Jan 01 Presented at: Global Communications Conference (GLOBECOM), IEEE; 2012; Anaheim, CA, USA. [doi: 10.1109/GLOCOM.2012.6503415]
39. Kennedy DO, Veasey RC, Watson AW, Dodd FL, Jones EK, Tiplady B, et al. Vitamins and psychological functioning: a mobile phone assessment of the effects of a B vitamin complex, vitamin C and minerals on cognitive performance and subjective mood and energy. Hum Psychopharmacol 2011;26(4-5):338-347. [doi: 10.1002/hup.1216] [Medline: 21751253]
40. Lachmann H, Fossum B, Johansson U, Karlgren K, Ponzer S. Promoting reflection by using contextual activity sampling: a study on students' interprofessional learning. J Interprof Care 2014 Sep;28(5):400-406. [doi: 10.3109/13561820.2014.907777] [Medline: 24754545]
41. Madan A, Cebrian M, Lazer D, Pentland A. Social sensing for epidemiological behavior change. Copenhagen, Denmark: ACM; 2010 Presented at: Proceedings of the 12th ACM international conference on Ubiquitous computing; September 26 - 29, 2010; Copenhagen, Denmark p. 291-300. [doi: 10.1145/1864349.1864394]
42. Muaremi A, Arnrich B, Tröster G. Towards measuring stress with smartphones and wearable devices during workday and sleep. Bionanoscience 2013;3:172-183 [FREE Full text] [doi: 10.1007/s12668-013-0089-2] [Medline: 25530929]
43. Muukkonen H, Hakkarainen K, Inkinen M, Lonka K, Salmela-Aro K. CASS-methods and tools for investigating higher education knowledge practices. 2008 Presented at: Proceedings of the 8th international conference on International conference for the learning sciences; 2008; Utrecht, The Netherlands p. 2.
44. Muukkonen H, Hakkarainen K, Kosonen K, Jalonen S, Heikkilä A, Lonka K, et al. Process-and context-sensitive research on academic knowledge practices: developing CASS-toolsmethods. New Brunswick, New Jersey, USA: International Society of the Learning Sciences; 2007 Presented at: Proceedings of the 8th iternational conference on Computer supported collaborative learning; 2007; Mahwah, NJ p. 541-543.

45. Ottaviani C, Medea B, Lonigro A, Tarvainen M, Couyoumdjian A. Cognitive rigidity is mirrored by autonomic inflexibility in daily life perseverative cognition. Biol Psychol 2015 Apr;107:24-30. [doi: 10.1016/j.biopsycho.2015.02.011] [Medline: 25749107]
46. Pärkkä J, Merilahti J, Mattila EM, Malm E, Antila K, Tuomisto MT, et al. Relationship of psychological and physiological variables in long-term self-monitored data during work ability rehabilitation program. IEEE Trans Inf Technol Biomed 2009 Mar;13(2):141-151. [doi: 10.1109/TITB.2008.2007078] [Medline: 19272856]
47. Pipingas A, Camfield DA, Stough C, Cox KH, Fogg E, Tiplady B, et al. The effects of multivitamin supplementation on mood and general well-being in healthy young adults. A laboratory and at-home mobile phone assessment. Appetite 2013 Oct;69:123-136. [doi: 10.1016/j.appet.2013.05.016] [Medline: 23727255]
48. Reitzel LR, Kendzor DE, Nguyen N, Regan SD, Okuyemi KS, Castro Y, et al. Shelter proximity and affect among homeless smokers making a quit attempt. Am J Health Behav 2014 Mar;38(2):161-169 [FREE Full text] [doi: 10.5993/AJHB.38.2.1] [Medline: 24629545]
49. Sano A, Phillips A, Yu A, McHill A, Taylor S, Jaques N, et al. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. 2015 Jan 01 Presented at: IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN); 2015; Cambridge, MA, USA. [doi: 10.1109/BSN.2015.7299420]
50. Sano A, Picard R. Stress Recognition Using Wearable Sensors and Mobile Phones. 2010 Jan 01 Presented at: Affective Computing and Intelligent Interaction (ACII), Humaine Association Conference on; 2013; Geneva, Switzerland. [doi: 10.1109/ACII.2013.117]
51. Sarker H, Sharmin M, Ali AA, Rahman MM, Bari R, Hossain SM, et al. Assessing the availability of users to engage in just-in-time intervention in the natural environment. Proc ACM Int Conf Ubiquitous Comput 2014;2014:909-920 [FREE Full text] [doi: 10.1145/2632048.2636082] [Medline: 25798455]
52. Vhaduri S, Ali A, Sharmin M, Hovsepian K, Kumar S. Estimating Drivers' Stress from GPS Traces. Proc Int Conf Automot User Interfaces Interact Veh Appl (2014) 2014 Sep 17;2014:909-920 [FREE Full text] [doi: 10.1145/2667317.2667335] [Medline: 25866847]
53. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. Studentlife: assessing mental health, academic performancebehavioral trends of college students using smartphones. 2014 Jan 01 Presented at: Proceedings of the ACM International Joint Conference on PervasiveUbiquitous Computing; September 13 - 17, 2014; New York, NY, USA p. 3-14. [doi: 10.1145/2632048.2632054]
54. Weppner J, Lukowicz P, Serino S, Cipresso P, Gaggioli A, Riva G. Smartphone based experience sampling of stress-related events. 2013 Presented at: Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare; 2013; Venice, Italy. [doi: 10.4108/icst.pervasivehealth.2013.252358]
55. Witkiewitz K, Desai SA, Steckler G, Jackson KM, Bowen S, Leigh BC, et al. Concurrent drinking and smoking among college students: an event-level analysis. Psychol Addict Behav 2012 Sep;26(3):649-654 [FREE Full text] [doi: 10.1037/a0025363] [Medline: 21895348]
56. Wray JM, Gray KM, McClure EA, Carpenter MJ, Tiffany ST, Saladin ME. Gender differences in responses to cues presented in the natural environment of cigarette smokers. Nicotine Tob Res 2015 Apr;17(4):438-442 [FREE Full text] [doi: 10.1093/ntr/ntu248] [Medline: 25762753]
57. Zenk SN, Horoi I, McDonald A, Corte C, Riley B, Odoms-Young AM. Ecological momentary assessment of environmental and personal factors and snack food intake in African American women. Appetite 2014 Dec;83:333-341. [doi: 10.1016/j.appet.2014.09.008] [Medline: 25239402]
58. Zenonos A, Khan A, Kalogridis G, Vatsikas S, Lewis T, Sooriyabandara M. HealthyOffice: Mood recognition at work using smartphones and wearable sensors. 2016 Jan 01 Presented at: IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops); March 14-18, 2016; Sydney, Australia. [doi: 10.1109/PERCOMW.2016.7457166]
59. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
60. Derogatis LR. The Derogatis Stress Profile (DSP): quantification of psychological stress. Adv Psychosom Med 1987;17:30-54. [Medline: 3591564]
61. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Appl Soc Psychol 2012;42(6):1320-1334. [doi: 10.1111/j.1559-1816.2012.00900.x]
62. Derogatis-tests. Synopsis of the Derogatis Stress Profile (DSP®) URL: http://www.derogatis-tests.com/dsp_synopsis.asp [accessed 2016-07-18] [WebCite Cache ID 6j5mGCp9S]
63. Wysocki J, Chemers MM, Rhodewalt F. Situational demand and self-reports of stress and illness: the moderating influence of self-monitoring. Basic Appl Soc Psych 2010 Jun 07;8(3):249-258. [doi: 10.1207/s15324834basp0803_5]

Abbreviations

ACM: Association for Computing Machinery EMA: ecological momentary assessment IEEE: Institute of Electrical and Electronics Engineers IT: information technology N/A: not available PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 26.07.16; peer-reviewed by V Osmani, F Beute, J Torous, M Egbring, A Ramachandran, TR Soron; comments to author 19.11.16; revised version received 14.12.16; accepted 18.12.16; published 13.02.17.

Please cite as:
Þórarinsdóttir H, Kessing LV, Faurholt-Jepsen M
Smartphone-Based Self-Assessment of Stress in Healthy Adult Individuals: A Systematic Review
J Med Internet Res 2017;19(2):e41
URL: http://www.jmir.org/2017/2/e41/
doi:10.2196/jmir.6397
PMID:28193600

Original Paper

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

Kati Anneli Kannisto[1,2], PhD; Joonas Korhonen[1,3], MHSc; Clive E Adams[4], MD; Marita Hannele Koivunen[1,2], PhD; Tero Vahlberg[5,6], MSc; Maritta Anneli Välimäki[1,6,7], RN, PhD

[1]Department of Nursing Science, University of Turku, Turku, Finland
[2]Satakunta Hospital District, Pori, Finland
[3]Turku University of Applied Sciences, Turku, Finland
[4]Institute of Mental Health, Division of Psychiatry, University of Nottingham, Nottingham, United Kingdom
[5]Department of Biostatistics, University of Turku, Turku, Finland
[6]Turku University Hospital, Turku, Finland
[7]Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Kati Anneli Kannisto, PhD
Department of Nursing Science
University of Turku
Joukahaisenkatu 3-5
Turku, FI-20014
Finland
Phone: 358 505312652
Fax: 358 23338400
Email: kaankan@utu.fi

Abstract

Background: Clinical trials are the gold standard of evidence-based practice. Still many papers inadequately report methodology in randomized controlled trials (RCTs), particularly for mHealth interventions for people with serious mental health problems. To ensure robust enough evidence, it is important to understand which study phases are the most vulnerable in the field of mental health care.

Objective: We mapped the recruitment and the trial follow-up periods of participants to provide a picture of the dropout predictors from a mHealth-based trial. As an example, we used a mHealth-based multicenter RCT, titled “Mobile.Net,” targeted at people with serious mental health problems.

Methods: Recruitment and follow-up processes of the Mobile.Net trial were monitored and analyzed. Recruitment outcomes were recorded as screened, eligible, consent not asked, refused, and enrolled. Patient engagement was recorded as follow-up outcomes: (1) attrition during short message service (SMS) text message intervention and (2) attrition during the 12-month follow-up period. Multiple regression analysis was used to identify which demographic factors were related to recruitment and retention.

Results: We recruited 1139 patients during a 15-month period. Of 11,530 people screened, 36.31% (n=4186) were eligible. This eligible group tended to be significantly younger (mean 39.2, SD 13.2 years, $P<.001$) and more often women (2103/4181, 50.30%) than those who were not eligible (age: mean 43.7, SD 14.6 years; women: 3633/6514, 55.78%). At the point when potential participants were asked to give consent, a further 2278 refused. Those who refused were a little older (mean 40.2, SD 13.9 years) than those who agreed to participate (mean 38.3, SD 12.5 years; t1842=3.2, $P<.001$). We measured the outcomes after 12 months of the SMS text message intervention. Attrition from the SMS text message intervention was 4.8% (27/563). The patient dropout rate after 12 months was 0.36% (4/1123), as discovered from the register data. In all, 3.12% (35/1123) of the participants withdrew from the trial. However, dropout rates from the patient survey (either by paper or telephone interview) were 52.45% (589/1123) and 27.8% (155/558), respectively. Almost all participants (536/563, 95.2%) tolerated the intervention, but those who discontinued were more often women (21/27, 78%; *P*=.009). Finally, participants' age (*P*<.001), gender (*P*<.001), vocational education (*P*=.04), and employment status (*P*<.001) seemed to predict their risk of dropping out from the postal survey.

Conclusions: Patient recruitment and engagement in the 12-month follow-up conducted with a postal survey were the most vulnerable phases in the SMS text message-based trial. People with serious mental health problems may need extra support during the recruitment process and in engaging them in SMS text message-based trials to ensure robust enough evidence for mental health care.

ClinicalTrial: International Standard Randomized Controlled Trial Number (ISRCTN): 27704027; http://www.isrctn.com/ISRCTN27704027 (Archived by WebCite at http://www.webcitation.org/6oHcU2SFp)

(J Med Internet Res 2017;19(2):e46) doi:10.2196/jmir.6417

KEYWORDS

text messaging; mobile health; antipsychotics; mental health; psychiatric services; methodological study

Introduction

Serious mental health problems are a major problem around the world [1]. They are associated with cognitive deficits, such as distortions in thinking or troubles in paying attention or working memory [2], a lack of treatment adherence [3], rehospitalization [4], and lifelong disability [5]. Mobile technology has become a popular way to deliver interventions to facilitate adherence to chronic disease management [6], including to people with serious mental health problems [7-9].

Recently, interventions with mobile phone technology (mHealth) have been applied to randomized controlled trials (RCTs) [10]. However, many methodological concerns in RCTs have been raised when mHealth interventions have targeted people with serious mental health problems [9]. Most concerns are related to inadequately reported details in the studies [10], such as the patient recruitment process [9], participant engagement in mHealth-based interventions [11], or technological details related to intervention delivery (eg, amount of undelivered text messages or patients changing a phone number) [12]. Because the mental health area continues to adopt new technologies in clinical practice [7,13,14], generating high-quality research is essential.

On the other hand, conducting mHealth-based research among people with mental health problems includes challenges [11]. First of all, reaching the target group may be difficult [11,15]. Patients may refuse to participate due to a fear of or suspicious thoughts about mobile devices [16]. They may also distrust the credibility of a mHealth intervention [11]. Other concerns related to mHealth use are user privacy, confidentiality, and online security [11]. People with serious mental health problems may be "digitally divided," which makes them vulnerable for not benefiting from mHealth services [17]. For example, not everyone is willing to use mobile technology [13], has access to technology [17,18], or is skilful or familiar with mobile phones [19]. Further, low engagement and discontinuation are fundamental problems in technology-based intervention studies [9,20]. The participants may stop the intervention because they feel that it is too complex, time consuming [11], or repetitive and, therefore, the intervention becomes a mere routine.

To better understand how mobile apps could be developed, evaluated, and implemented into routine care, it is important to truly understand which study phases make the RCT the most vulnerable in the field of mental health care. Still, many important parts of the study methodology are inadequately reported in RCTs, particularly regarding interventions targeting people with serious mental health problems [9]. Therefore, we mapped the recruitment and the 12-month trial follow-up periods in order to provide a picture of the dropout predictors from a mHealth-based multicenter RCT, titled "Mobile.Net" (ISRCTN: 27704027). Recruitment and engaging participants in trials involving psychosis is problematic; there are numerous ways in which it can go wrong (eg, consenting, attrition). This paper describes a case study of recruitment and follow-up processes, and problems in this context, based on the Mobile.Net trial. Multiple regression was used to identify which demographic factors are related to recruitment and retention. The main results of the Mobile.Net trial will be reported elsewhere.

Methods

Mobile.Net Trial

Mobile.Net is a nationwide multicenter randomized controlled two-armed trial. The Mobile.Net trial evaluated the effects of tailored short message service (SMS) text messages constructed to encourage patient medication adherence and outpatient care for adult patients with psychosis [21]. Participants in the intervention group received semiautomatic text messages for 12 months (approximately 10 per month; range 2-25 text messages) based on their preferences [22]. They were able to decide the amount, timing, and frequency of the SMS text messages delivered. They were also able to change the content or timing of the messages during the trial. Treatment as usual was offered to all participants. The study was carried out according to the Declaration of Helsinki and approved by the Ethics Committee of the Hospital District of Southwest Finland. Written informed consent was obtained from the study participants after they were given a complete description of the study.

Population

There were a total of 1139 participants, men and women, ranging in age from 18 to 65 years. Each participant had a continuing prescription for antipsychotic medication, access to a mobile phone, and the ability to use the Finnish language. After participants were recruited, they were then randomized. Forensic patients and those having a planned nonacute treatment period were excluded from the study [21].

Procedure

Recruitment, including activities conducted before and during participant enrollment [23], occurred face-to-face in 45 psychiatric hospital wards in Finland (between September 5, 2011 and November 30, 2012). Research nurses in each study ward performed chart reviews to check the eligibility of each patient admitted to the study ward [24]. After completion of the baseline data, patients were allocated randomly (computer-based randomization with four block randomization) into two groups (SMS text message intervention group or control group) [21].

Attrition, including actions after enrollment in the study [23], was assessed at the 12-month follow-up period (between September 5, 2012 and December 31, 2013). Attrition was defined as the loss of eligible participants from the study groups [25]. To gather follow-up data, a postal survey, including the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q) by Endicott et al [26] and the Client Satisfaction Questionnaire (CSQ-8) by Atkisson and Greenfield [27], was carried out. Follow-up data were collected by members of the research group. During the study period and follow-up, each point of contact with participants was completely tracked and recorded. Attrition was assessed during the intervention and during the follow-up.

Measures

Data for this paper were divided into two categories: data concerning patient recruitment and data relating to attrition (Textbox 1).

Textbox 1. Measures concerning patient recruitment and attrition from the study.

Recruitment (variable: measurement)

1) Screened: (n)

2) Eligible: n (%)

3) Consent not asked: n (%)

4) Refused: n (%)

5) Enrolled: n (%)

Recruitment speed: n/day

Attrition (variable: measurement)

1) During SMS intervention: dropout rate

2) During follow-up

- Telephone interview: dropout rate
- Participant’s notification: dropout rate
- Postal survey: dropout rate
- Register data retrieval: dropout rate

Recruitment data were categorized into five groups: (1) patients screened for eligibility, (2) eligible participants, (3) eligible participants whose consent was not requested, (4) participants who refused to participate at the point of contact in the psychiatric ward, and (5) those who consented to participate. Outcomes were recorded as screened, eligible, consent not asked, refused, and enrolled. In addition, to track the pace of recruitment, a record of all identified, screened, eligible, unwilling, and successfully recruited patients was kept using a specific monitoring sheet developed for the trial. Patient flow was monitored and recorded daily on the study wards. Daily progress of patient recruitment was reported as “recruitment pace” (ie, how many new patients were recruited e each day) [28].

Attrition data were categorized into two groups: (1) attrition during SMS text message intervention and (2) attrition during the 12-month follow-up period [29]. Attrition during the follow-up was then divided into four categories, based on the follow-up data collection method: (1) telephone interview, (2) participants’ notification (ie, withdrew from the follow-up survey), (3) postal survey, or (4) register data retrieval from the Finnish National Care Register for Health Care [30]. Outcomes were reported as dropout rates. In our study, the dropout rates were calculated based on telephone or postal survey responses, or by whose data were not available in register data retrieval due to an incorrectly entered ID.

Statistical Analysis

Descriptive statistics (frequency, percentage, mean, standard deviation) were used to describe participants’ demographic characteristics, recruitment, and attrition metrics (study participants lost in the follow-up). The demographic variables examined included age, gender, marital status (lives alone, ie, single, divorced, or widowed; lives with someone, ie, married), vocational education (none, vocational education), employment status (employed/self-employed, retired, student, job seeker), diagnosis (*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* [31]), and age at first contact with psychiatric services. To analyze possible differences between patients who participated in the study and those who dropped out, *t* tests and chi-square tests were used.

Multiple logistic regression analysis was used to determine predictors of dropping out of the 12-month postal survey follow-up. Participants’ demographic characteristics, including age, gender, marital status, vocational education, and employment status, were chosen as predictors and added to the analysis [8,32]. All data were analyzed using SPSS version 21. A *P* value <.05 was interpreted as a statistically significant difference.

Results

Recruitment

A total of 11,530 patients admitted within psychiatric inpatient hospital wards were screened during the 15-month (453 days) recruitment period. There were 6565 who did not meet the eligibility criteria. A total of 779 patients dropped out before the eligibility assessment because they were transferred to another ward or rapidly discharged from hospital. Of the candidates who were screened, 36.31% (4186/11,530) appeared eligible.

Of the 4186 eligible patients, informed consent was asked from 3417 (81.63%). Informed consent was not asked in 18.37% (769/4186) of the cases because the person was quickly discharged from the ward, absconded from hospital, or the research nurses simply forgot to ask.

When age and gender of the screened noneligible and eligible patients were compared, it was found that the eligible patients were generally younger than the noneligible patients (*P*<.001). Men were more often noneligible than women (*P*<.001) (Table 1).

Table 1. Demographic characteristics comparable across all stages.

Stage of study	N	Age (years)				Gender (male)		
		Mean (SD)	Range	*t* (df)	*P*	n/N (%)	χ^2_1	*P*
Screened	11,530	41.1 (14.6)	16-90	–9.86 (3492)	<.001	6164/11,461 (53.78)		
Noneligible	6565	43.7 (16.1)	16-90			3633/6514 (55.77)	30.7	<.001
Eligible	4186	39.2 (13.2)	18-65			2103/4181 (50.30)		
Point of consent				3.23 (1842)	.001		0.3	.59
Refused	2278	40.2 (13.9)	18-65			1142/2274 (50.22)		
Randomized	1139	38.3 (12.5)	18-65			560/1139 (49.17)		
Intervention period				–0.73 (561)	.47		7.2	.009
Completers	536	38.5 (12.7)	18-65			261/536 (48.7)		
Dropouts	27	40.3 (13.0)	21-63			6/27 (22.2)		
Follow-up period				–1.28 (36)	.21		10.1	.002
Completers	1088	38.3 (12.5)	18-65			545/1088 (50.09)		
Withdrawals	35	41.1 (12.6)	18-63			8/35 (22.9)		
Postal survey				–8.14 (1120)	<.001		18.5	<.001
Completers	534	41.5 (12.6)	18-65			227/534 (41.0)		
Dropouts	589	35.5 (11.8)	18-65			326/589 (59.0)		

Out of the 3417 eligible participants whose consent was asked, 2278 patients (66.67%) refused to participate in the study. Although reasons for refusal were not asked due to ethical guideline requirements [33], some patients voluntarily offered explanations, such as they did not know how to use mobile phones or text messages, their mobile phone was broken, or they did not have a mobile subscription at the that time. Patients who refused to participate were older than consenting, randomized patients (*P*=.001) (Table 1).

The pace of recruitment was analyzed based on the number of new patients recruited each day. At the beginning of the study, recruitment was slow. The recruitment rate reached its peak 15 months after enrollment started. For every 10 screens completed, one person was successfully enrolled, at an average recruitment speed of 76 participants each month (2.5 participants per day).

Of the 1139 patients who were enrolled in the study, the data of 16 participants were excluded due to either the withdrawal of informed consent (n=10), the patient did not meet the inclusion criteria (n=5), or a recruitment error (n=1). This left us with a total of 1123 participants (intervention group: n=563; control group: n=560).

Attrition

Attrition During SMS Text Message Intervention Period

A total of 569 eligible participants were allocated to a group to receive tailored SMS text messages for 12 months. The data of six participants were excluded from the analyses due to either a lack of written informed consent (n=2), the patient did not meet the inclusion criteria (n=3), or an erroneous randomization to study group (n=1). This left us 563 participants.

Of the 563 participants who received text messages, 27 dropped out during the 12-month intervention period (4.8%). In cases where a patient did not want to continue with the text message intervention, the researchers were notified by the participant, a relative, or a research nurse. Three participants dropped out before the intervention even began, and 24 within the 12-month intervention period [22]. Participants who dropped out during the intervention were still included in the study [34]. We observed that intervention dropouts were more often women than men (χ^2_1=7.2, P=.009), but found no other statistically significant differences (Multimedia Appendix 1).

Attrition During the 12-Month Follow-Up Period

Information about participants who dropped out after the 12-month follow-up was divided into four categories based on the data collection method: (1) telephone interview, (2) participants' notification (ie, withdrew from the follow-up survey), (3) postal survey, or (4) register data retrieval.

First, telephone interviews (for the intervention group only) were conducted after the 12-month text message intervention to explore participants' feedback on the text message service (n=569). We attempted to reach 558 participants by telephone for an interview; after the telephone calls were made, we had 403 completed questionnaires (response rate 72.2%, 403/558) [35]. The dropout rate from these telephone interviews was 27.8% (155/558). Dropouts were younger, usually men, without a vocational education, and were also younger at the time of first contact with psychiatric services compared with those who completed the questionnaire [35].

Second, 35 participants expressed that they wanted to withdraw from the follow-up surveys (intervention group: 5.5%, 31/563; control group: 0.7%, 4/560; χ^2_1=21.4, P<.001). Follow-up surveys were not conducted with these participants, but their register data were retrieved. The dropout rate regarding follow-up surveys was 3.12% (35/1123). Women requested to withdraw from the follow-up surveys more often than men did (χ^2_1=10.1, P=.002). We found no other statistically significant differences (Multimedia Appendix 1).

Third, a postal survey (n=1123) was conducted after the 12-month study period to measure participants' quality of life (Q-LES-Q [26]) and satisfaction with the treatment (CSQ-8 [27]). Of 1123 participants, 589 did not return the postal survey (dropout rate: 52.45%, 589/1123). There was a statistically significant difference in demographic characteristics between postal survey completers (ie, returned the completed postal survey) and dropouts (Multimedia Appendix 1).

Fourth, register data retrieval was conducted after the 12-month follow-up period. Out of 1123 participants, the register data of four participants were not available from the Finnish National Care Register for Health Care [30], making the dropout rate 0.36%. Demographic characteristics of participants whose register data were available will be reported elsewhere.

Factors Related to Patient Attrition

Through a logistic regression analysis, Table 2 illustrates the associations between participants' demographic characteristics (age, gender, marital status, vocational education, and employment status) and risk of dropping out from the postal survey. Odds ratios were not estimated from other variables due to missing data. Participants' age, gender, vocational education, and employment status seemed to predict their risk of dropping out from the postal survey. The participants in this group were older, more often women, had a vocational education, and were more often retired.

Table 2. Associations between participants' demographic characteristics and risk of dropping out of the postal survey (N=1123).

Demographic characteristics	OR (95% CI)	*P*
Age	0.96 (0.95-0.97)	<.001
Gender		
Female	1	<.001
Male	1.63 (1.27-2.11)	
Marital status		
Lives with someone	1	.46
Lives alone	1.12 (0.83-1.50)	
Vocational education		
Vocational education	1	.04
None	1.37 (1.01-1.84)	
Employment status		
Student	1	.001
Employed/self-employed	1.65 (1.01-2.70)	.045
Retired	2.29 (1.45-3.61)	<.001
Job seeker	2.44 (1.50-3.97)	<.001

Discussion

Principal Results

The results of this study demonstrate that it was challenging to recruit, and engage, participants in an SMS text message-based trial follow-up. One-third of patients (36.31%) appeared eligible, and two-thirds of eligible patients (66.67%) refused to participate. Participants were well engaged with the SMS text message intervention provided, but their engagement with the trial follow-up varied: the highest being with the register data retrieval (99.64%) and lowest with the postal survey (47.55%). Participants' demographic characteristics (age, gender, vocational education, and employment status) were seen as dropout predictors.

In our study, within the context of psychiatric inpatient care, we were able to recruit 1139 individuals (33.33%) out of 3417 eligible participants, whose consent was requested. Age and gender tended to be factors influencing recruitment and refusal. Our refusal rate was 66.67% (2278/3417), which is in line with previous studies suggesting that high refusal rates are a major problem faced during the recruitment process [11,15]. Patients' illness-related issues [14,19] and perceived stigma were seen as barriers to participation in previous studies [36,37]. This may also be the case in our study, which focused on people treated with antipsychotic medication recruited from hospital wards providing psychiatric care.

Lack of interest in the trial [11], lack of motivation to use mobile technology-based interventions [13], or a lack of capacities may also have affected the refusal rates in our trial. During the participant screening, some patients expressed their inability to use mobile phones or text messaging, stating that as the reason for their refusal. This was unexpected because mobile phone text messages have been proven to be feasible and acceptable among people with severe mental health problems [38], as well as easy to use [35]. Therefore, our findings may reveal that a digital divide may still exist in the area of mental health, although relevant literature identifies this concept as promising [14]. Previous literature also shows that people with mental health problems are able and willing to use mobile-based interventions when given the opportunity [38,39]. In our case, it might have been useful to train those who said they did not know how to use mobile phones, and to offer mobile phones to those who did not have them. This might have given us important knowledge about the digital divide, especially those who are not so familiar with the mobile technology.

The attrition rate during the SMS text message intervention was low (4.8%). This finding does not support previous findings stating that low engagement and discontinuation are major problems in intervention studies [11,14]. However, it is possible that participants who stopped using the SMS text message intervention did not notify the researchers. What we found here was that women dropped out from the SMS text message intervention more often than men did. This is contradictory to the results of Ben-Zeev and colleagues [8], who found that women were significantly more engaged with mHealth interventions than men were. Further, participants' demographic characteristics (age, gender, vocational education, and employment status) predicted the risk of leaving the study early and, subsequently, not participating in the follow-up. This is in line with previous studies, which reported that participants who left the studies early differed from those who were retained [8,32]. Therefore, it is important to identify feasible and useful mHealth interventions targeting different patient groups and, further, to identify factors that facilitate or prevent patient engagement with mHealth interventions.

Patient engagement in the trial during the follow-up varied depending on the source of data collected, the highest being in the register data retrieval (99.64%) and the lowest in the postal survey (47.55%). Low engagement and discontinuation have been found as fundamental problems in technology-based intervention studies [9,20]. Participants' age, gender, vocational education, and employment status seemed to predict their risk of dropping out from the postal survey. This group of older, women, and retired participants may also reveal the digital divide in this patient group. It may be that they just do not value technology as much as younger people do and, therefore, did not participate in the postal survey. Recently, however, technology usage has increased among older adults [40]. This is promising and may give researchers new hints on how to engage participants in trials. Given this, it is important to consider which data collection methods are appropriate to use among people with serious mental health problems, and to identify factors that facilitate patient engagement in mHealth-based trial follow-ups.

Limitations

Our study has some limitations. The recruitment data concerning information about screened patients lacked some information, especially about patients' ages. Therefore, results related to patient demographics concerning recruitment have to be handled with caution. More importantly, we did not gather knowledge about the participants' actual SMS text message use. Therefore, we lack knowledge about participants' true engagement with the SMS text message intervention. However, according to our findings before the study actually started, participants were very satisfied with the intervention [35].

A key strength of this study was in its large nationwide sample of people treated with antipsychotic medication. Another was that, to the best of our knowledge, this was the largest trial evaluating a text message system. Our findings regarding attrition are important for those conducting similar RCTs among people with severe mental health problems, although this group may well have different issues with the technology when compared with others [17,41].

Conclusions

Initial patient recruitment and then engagement in the 12-month postal survey follow-up were the most vulnerable phases in the SMS text message-based trial. This may indicate that people with serious mental health problems may need extra support during the recruitment process, and necessitate further support to engage in completion of these follow-up questionnaires—at least within SMS text message trials.

Researchers should acknowledge the possible digital divide for people with serious mental health problems, and choose convenient and efficient data collection methods for study follow-ups. At follow-up, for Mobile.Net, high-grade routine data were almost complete. Methods of trials should take much more consideration of the nature of the target group of participants; otherwise, evidence is dogged with high attrition with the accompanying speculation of researchers. No statistical technique or learned speculation can make up for loss to follow-up. The solutions are likely to vary for different client groups. We think more research is needed both to investigate the support of the recruitment process and methods of follow-up in technology-based RCTs. Asking people to complete forms that are likely to result in grossly incomplete datasets could be considered an unethical—and potentially dangerous—waste of time and resources.

Acknowledgments

The authors would like to warmly thank all the institutions who awarded grants for Professor Välimäki: the Academy of Finland (132581), Turku University Hospital (EVO 13893), Satakunta Hospital District (EVO 12/2010, 81096), Foundations' Professor Pool, the Finnish Cultural Foundation, and the University of Turku. Many thanks also go to the patients participating in the Mobile.Net study, research nurses, and the staff of the research organizations and wards, without whom the realization of this study would not have been possible. The authors especially wish to thank Kaisa Kauppi, PhD, Minna Anttila, PhD, and Sanna Suni, MA, for their valuable help with monitoring, data collection, and data entry. They also thank the Mobile.Net Safety Committee Group for their efforts and support throughout this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional demographic characteristics comparable across latter stages of study.

[PDF File (Adobe PDF File), 70KB - jmir_v19i2e46_app1.pdf]

References

1. Word Health Organization. Schizophrenia URL: http://www.who.int/mental_health/management/schizophrenia/en/ [accessed 2016-11-01] [WebCite Cache ID 6lhG6roQ8]
2. Harvey PD, Strassnig M. Predicting the severity of everyday functional disability in people with schizophrenia: cognitive deficits, functional capacity, symptoms, and health status. World Psychiatry 2012 Jun;11(2):73-79 [FREE Full text] [Medline: 22654932]
3. Velligan DI, Weiden PJ, Sajatovic M, Scott J, Carpenter D, Ross R, Expert Consensus Panel on Adherence Problems in Serious and Persistent Mental Illness. The expert consensus guideline series: adherence problems in patients with serious and persistent mental illness. J Clin Psychiatry 2009;70 Suppl 4:1-46; quiz 47 [FREE Full text] [Medline: 19686636]
4. Haddad PM, Brain C, Scott J. Nonadherence with antipsychotic medication in schizophrenia: challenges and management strategies. Patient Relat Outcome Meas 2014;5:43-62 [FREE Full text] [doi: 10.2147/PROM.S42735] [Medline: 25061342]
5. Świtaj P, Anczewska M, Chrostek A, Sabariego C, Cieza A, Bickenbach J, et al. Disability and schizophrenia: a systematic review of experienced psychosocial difficulties. BMC Psychiatry 2012 Nov 09;12:193 [FREE Full text] [doi: 10.1186/1471-244X-12-193] [Medline: 23137171]
6. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
7. Gaebel W, Großimlinghaus I, Kerst A, Cohen Y, Hinsche-Böckenholt A, Johnson B, et al. European Psychiatric Association (EPA) guidance on the quality of eMental health interventions in the treatment of psychotic disorders. Eur Arch Psychiatry Clin Neurosci 2016 Mar;266(2):125-137. [doi: 10.1007/s00406-016-0677-6] [Medline: 26874958]
8. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for schizophrenia: patient engagement with a mobile phone intervention following hospital discharge. JMIR Ment Health 2016;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
9. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. J Med Internet Res 2016 Jun 10;18(6):e135 [FREE Full text] [doi: 10.2196/jmir.5066] [Medline: 27287668]
10. 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. BMJ 2016;352:i1174. [Medline: 26988021]
11. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015;6:27882 [FREE Full text] [Medline: 25994025]

12. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]
13. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
14. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of “mHealth” among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2015 Sep 22;42(2):448-455. [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
15. Furimsky I, Cheung AH, Dewa CS, Zipursky RB. Strategies to enhance patient recruitment and retention in research involving patients with a first episode of mental illness. Contemp Clin Trials 2008 Nov;29(6):862-866. [doi: 10.1016/j.cct.2008.07.005] [Medline: 18721902]
16. Depp CA, Harmell AL, Vahia IV, Mausbach BT. Neurocognitive and functional correlates of mobile phone use in middle-aged and older patients with schizophrenia. Aging Ment Health 2016;20(1):29-35 [FREE Full text] [doi: 10.1080/13607863.2015.1008987] [Medline: 25768842]
17. Ennis L, Rose D, Denis M, Pandit N, Wykes T. Can't surf, won't surf: the digital divide in mental health. J Ment Health 2012 Aug;21(4):395-403 [FREE Full text] [doi: 10.3109/09638237.2012.689437] [Medline: 22712756]
18. Harvey PD, Keefe RSE. Technology, society, and mental illness: challenges and opportunities for assessment and treatment. Innov Clin Neurosci 2012 Nov;9(11-12):47-50 [FREE Full text] [Medline: 23346519]
19. Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC. Strategies for mHealth research: lessons from 3 mobile intervention studies. Adm Policy Ment Health 2015 Mar;42(2):157-167. [doi: 10.1007/s10488-014-0556-2] [Medline: 24824311]
20. 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]
21. Välimäki M, Hätönen H, Adams CE. Mobile.net: mobile telephone text messages to encourage adherence to medication and to follow up with people with psychosis: methods and protocol for a multicenter randomized controlled two-armed trial. JMIR Res Protoc 2012;1(2):e8 [FREE Full text] [doi: 10.2196/resprot.2136] [Medline: 23611874]
22. Kauppi K, Kannisto KA, Hätönen H, Anttila M, Löyttyniemi E, Adams CE, et al. Mobile phone text message reminders: Measuring preferences of people with antipsychotic medication. Schizophr Res 2015 Oct;168(1-2):514-522. [doi: 10.1016/j.schres.2015.07.044] [Medline: 26293215]
23. Blanton S, Morris DM, Prettyman MG, McCulloch K, Redmond S, Light KE, et al. Lessons learned in participant recruitment and retention: the EXCITE trial. Phys Ther 2006 Nov;86(11):1520-1533. [doi: 10.2522/ptj.20060091] [Medline: 17079752]
24. Kim R, Hickman N, Gali K, Orozco N, Prochaska JJ. Maximizing retention with high risk participants in a clinical trial. Am J Health Promot 2014;28(4):268-274. [doi: 10.4278/ajhp.120720-QUAN-355] [Medline: 23875989]
25. Amico KR. Percent total attrition: a poor metric for study rigor in hosted intervention designs. Am J Public Health 2009 Sep;99(9):1567-1575. [doi: 10.2105/AJPH.2008.134767] [Medline: 19608965]
26. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. Psychopharmacol Bull 1993;29(2):321-326. [Medline: 8290681]
27. Attkisson CC, Greenfield TK. The UCSF Client Satisfaction Scales: I. The Client Satisfaction Questionnaire-8. In: Maruish ME, editor. The Use Of Psychological Testing For Treatment Planning And Outcomes Assessment: Volume 1: General Considerations (3rd Edition). Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
28. Giguere AM, Labrecque M, Légaré F, Grad R, Cauchon M, Greenway M, et al. Feasibility of a randomized controlled trial to evaluate the impact of decision boxes on shared decision-making processes. BMC Med Inform Decis Mak 2015;15:13. [doi: 10.1186/s12911-015-0134-x] [Medline: 25880757]
29. Karlson CW, Rapoff MA. Attrition in randomized controlled trials for pediatric chronic conditions. J Pediatr Psychol 2009 Aug;34(7):782-793 [FREE Full text] [doi: 10.1093/jpepsy/jsn122] [Medline: 19064607]
30. National Institute for Health and Welfare. 2014. Care register for health care URL: https://www.thl.fi/en/web/thlfi-en/statistics/information-on-statistics/register-descriptions/care-register-for-health-care [accessed 2016-07-26] [WebCite Cache ID 6jIJsCae5]
31. The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. Geneva: World Health Organization; 1993. URL: http://www.who.int/classifications/icd/en/GRNBOOK.pdf [accessed 2016-07-26] [WebCite Cache ID 6jIJxmaea]
32. Villeneuve K, Potvin S, Lesage A, Nicole L. Meta-analysis of rates of drop-out from psychosocial treatment among persons with schizophrenia spectrum disorder. Schizophr Res 2010 Aug;121(1-3):266-270. [doi: 10.1016/j.schres.2010.04.003] [Medline: 20452749]
33. Vargas-Barón E. Planning Policies for Early Childhood Development: Guidelines for Action. Paris: United Nations Educational, Scientific and Cultural Organisation (UNESCO); 2005. URL: http://www.unicef.org/ceecis/0611-ECDPolicyGuidelines_en.pdf [accessed 2016-07-26] [WebCite Cache ID 6jIK2c35d]
34. Shadish W, Cook T, Campbell D. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin; 2002.

35. Kannisto KA, Adams CE, Koivunen M, Katajisto J, Välimäki M. Feedback on SMS reminders to encourage adherence among patients taking antipsychotic medication: a cross-sectional survey nested within a randomised trial. BMJ Open 2015;5(11):e008574 [FREE Full text] [doi: 10.1136/bmjopen-2015-008574] [Medline: 26553830]
36. Jørgensen R, Munk-Jørgensen P, Lysaker PH, Buck KD, Hansson L, Zoffmann V. Overcoming recruitment barriers revealed high readiness to participate and low dropout rate among people with schizophrenia in a randomized controlled trial testing the effect of a Guided Self-Determination intervention. BMC Psychiatry 2014 Feb 03;14:28 [FREE Full text] [doi: 10.1186/1471-244X-14-28] [Medline: 24490977]
37. Woodall A, Morgan C, Sloan C, Howard L. Barriers to participation in mental health research: are there specific gender, ethnicity and age related barriers? BMC Psychiatry 2010;10:103 [FREE Full text] [doi: 10.1186/1471-244X-10-103] [Medline: 21126334]
38. Ben-Zeev D, Kaiser SM, Krzos I. Remote “hovering” with individuals with psychotic disorders and substance use: feasibility, engagement, and therapeutic alliance with a text-messaging mobile interventionist. J Dual Diagn 2014;10(4):197-203 [FREE Full text] [doi: 10.1080/15504263.2014.962336] [Medline: 25391277]
39. Ben-Zeev D. Mobile technologies in the study, assessment, and treatment of schizophrenia. Schizophr Bull 2012 May;38(3):384-385 [FREE Full text] [doi: 10.1093/schbul/sbr179] [Medline: 22234487]
40. Smith A. Pew Research Center. 2014 Apr 03. Older adults and technology use URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2016-11-01] [WebCite Cache ID 6lhK47uHh]
41. McCauley-Elsom K, Gurvich C, Lee S, Elsom S, O'Connor M, Kulkarni J. Vulnerable populations and multicentred research. Int J Ment Health Nurs 2009 Apr;18(2):108-115. [doi: 10.1111/j.1447-0349.2008.00590.x] [Medline: 19290974]

Abbreviations

CSQ-8: Client Satisfaction Questionnaire
Q-LES-Q: Quality of Life Enjoyment and Satisfaction Questionnaire
RCT: randomized controlled trial
SMS: short message service

Edited by G Eysenbach; submitted 27.07.16; peer-reviewed by T Wykes, D Robotham; comments to author 22.09.16; revised version received 02.11.16; accepted 05.12.16; published 21.02.17.

Please cite as:
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;19(2):e46
URL: http://www.jmir.org/2017/2/e46/
doi:10.2196/jmir.6417
PMID:28223262

Original Paper

Scientific Misconduct and Social Media: Role of Twitter in the Stimulus Triggered Acquisition of Pluripotency Cells Scandal

Yuya Sugawara[1,2,3], MS; Tetsuya Tanimoto[4], MD; Shoko Miyagawa[5], MSHI, PhD; Masayasu Murakami[2], BEC; Atsushi Tsuya[1], PhD; Atsushi Tanaka[6], PhD; Masahiro Kami[7], MD, PhD; Hiroto Narimatsu[3,8], MD, PhD

[1]Institute for Promotion of Medical Science Research, Yamagata University, Faculty of Medicine, Yamagata, Japan
[2]Department of Health Policy Science, Graduate School of Medical Science, Yamagata University, Yamagata, Japan
[3]Cancer Prevention and Control Division, Kanagawa Cancer Center Research Institute, Yokohama, Japan
[4]Navitas Clinic, Tokyo, Japan
[5]Faculty of Nursing and Medical Care, Keio University, Fujisawa, Japan
[6]Graduate School of Science and Engineering, Yamagata University, Yonezawa, Japan
[7]Medical Governance Research Institute, Tokyo, Japan
[8]Department of Public Health, Graduate School of Medical Science, Yamagata University, Yamagata, Japan

Corresponding Author:
Hiroto Narimatsu, MD, PhD
Cancer Prevention and Control Division
Kanagawa Cancer Center Research Institute
2-3-2 Nakao
Asahi-ku
Yokohama, 241-8515
Japan
Phone: 81 45 520 2222
Fax: 81 45 520 2216
Email: hiroto-narimatsu@umin.org

Abstract

Background: The academic scandal on a study on stimulus-triggered acquisition of pluripotency (STAP) cells in Japan in 2014 involved suspicions of scientific misconduct by the lead author of the study after the paper had been reviewed on a peer-review website. This study investigated the discussions on STAP cells on Twitter and content of newspaper articles in an attempt to assess the role of social compared with traditional media in scientific peer review.

Objective: This study examined Twitter utilization in scientific peer review on STAP cells misconduct.

Methods: Searches for tweets and newspaper articles containing the term “STAP cells” were carried out through Twitter’s search engine and Nikkei Telecom database, respectively. The search period was from January 1 to July 1, 2014. The nouns appearing in the “top tweets” and newspaper articles were extracted through a morphological analysis, and their frequency of appearance and changes over time were investigated.

Results: The total numbers of top tweets and newspaper articles containing the term were 134,958 and 1646, respectively. Negative words concerning STAP cells began to appear on Twitter by February 9-15, 2014, or 3 weeks after Obokata presented a paper on STAP cells. The number of negative words in newspaper articles gradually increased beginning in the week of March 12-18, 2014. A total of 1000 tweets were randomly selected, and they were found to contain STAP-related opinions (43.3%, 433/1000), links to news sites and other sources (41.4%, 414/1000), false scientific or medical claims (8.9%, 89/1000), and topics unrelated to STAP (6.4%, 64/1000).

Conclusions: The discussion on scientific misconduct during the STAP cells scandal took place at an earlier stage on Twitter than in newspapers, a traditional medium.

(J Med Internet Res 2017;19(2):e57) doi:10.2196/jmir.6706

KEYWORDS

web 2.0; bioethics; Internet; mass media

Introduction

In recent years, the number of cases of scientific misconduct, including fabrication, falsification, and plagiarism has increased [1]. Misconduct damages scientific progress and public trust; in addition, the ensuing incorrect research results threaten people’s health [2]. As of May 2012, 2047 papers in the fields of medicine, biology, and life science were retracted from PubMed, a fully accessible database on biomedical literature. The reasons for the retraction were error (21.3%), fraud or suspected fraud (43.4%), duplicate publication (14.2%), and plagiarism (9.8%) [3,4]. Japan had the third largest number of retracted papers due to fraud or suspected fraud in the world [5].

The three major historical scandals of scientific misconduct were German physicist Jan Hendrik Schön’s fraudulent superconductor breakthroughs at Bell Labs in the United States in 2002, South Korean researcher Hwang Useok’s fabrication of embryonic stem cells in 2005, and Japanese stem-cell biologist Haruko Obokata’s and her fellow researchers’ claims on stimulus-triggered acquisition of pluripotency (STAP) cells in 2014 [6,7]. Suspicions on the three studies were raised in different arenas. Researchers who were unable to replicate Schön’s results raised their concerns in a conventional reseacher’s community [8], whereas discussions and debates on Useok’s study occurred on Korean and Japanese Web-based message boards, respectively [9]. The allegations of misconduct of Obokata, who led the STAP cells study, spread on Twitter after their paper had been publicly reviewed on PubPeer [10,11]. The hurling of accusations worldwide on a single paper by a large number of Twitter users, including many nonspecialist members of the public, attracted widespread attention. Their misconducts included “copying and pasting,” which were familiar methods to the public; thus, the public could join the discussion.

The STAP cells scandal demonstrated how Twitter enables the rapid spread of information through sharing between multiple users, allowing numerous users to obtain the information simultaneously. Suspicions on research papers, thus, can be raised on Twitter by multiple people, making the social media site a useful tool for discussion and debate. In fact, previous studies have found that researchers engage in discussions on their studies via Twitter [12,13]. Several studies have examined the concern expressed by the chief editor of a scientific journal regarding the self-plagiarism conducted by a chemist [14] and the controversy regarding Felisa Wolfe-Simon’s claims on Twitter about the bacteria that lived without phosphorous [15]. Through Twitter, a rapid response to questions on misconduct through a debate is possible and discussions lead to an exchange of a diverse range of opinions. Such processes generate more questions and play a role in dispute resolution. However, a widespread controversy on Twitter may have a restraining effect on the concerned researchers, and Twitter’s roles as a tool for discussion and later dispute resolution are yet to be determined.

This study investigated discussions on STAP cells on Twitter and the content of newspaper articles in an attempt to differentiate social from traditional media. To identify changes in the tone of Twitter and newspapers over time, collected tweets and newspaper articles on STAP were decomposed through a morphological analysis. Furthermore, to ascertain social media’s role in resolving scientific misconduct, Twitter’s role in the STAP cells scandal was compared with that of newspapers, which represent the traditional media.

Methods

Collection of HTML Files

Searches for tweets and newspaper articles containing the term “STAP cells” were carried out for 6 months (26 weeks) from January 1 to July 1, 2014. Tweets were extracted through Twitter’s search engine. A top tweets search was performed every week; top tweets are tweets that have been retweeted or replied to by several users and selected through an algorithm developed by Twitter [16]. The newspaper article search covered Japan’s 5 major national newspapers (*Asahi Shimbun*, *Yomiuri Shimbun*, *Mainichi Shimbun*, *Sankei Shimbun*, and *Nihon Keizai Shimbun*). The 5 major national newspapers published a combined total of 23,543 thousand copies every half-year on average [17,18]. Articles were extracted from Nikkei Telecom, a Japanese newspaper article database [19].

Tweets and newspaper articles generated from the search were saved as HTML files by using the Web browser function. One HTML file contained a week worth of search, except from April 9 to 15, when the number of tweets was so high that weekly data could not be saved in one file. As such, HTML data were saved separately for each day. The syntax for designating cut-off dates provided on each of the official sites was used in each search.

Extraction of Japanese Text Data

Text data were extracted by stripping the HTML tags from the HTML files saved from the Web browser. The dates, account names, text of the tweets, and the titles and text of newspaper articles were extracted from the Twitter HTML and newspaper article files, respectively. Nadeshiko (a Japanese programming software, Free edition, Kujirahando, Japan) was used to write a program that would eliminate tags and extract the text.

Text Processing

To generate the relevant words for the morphological analysis, the dates, account names, URLs, and graphic characters were eliminated from the tweet text data. The following frequently used terms were also eliminated from both tweets and newspaper articles: STAP, stap, cells, Obokata, Haruko, RIKEN Research Center, and RIKEN. The free bulk text processing software Text Search and Substitute.NET (TextSS.NET) version 5.21 (Yamashita-Y, Japan) was used for text processing.

Morphological Analysis

Morphological analysis is one of the basic techniques used in Japanese text mining analysis. It is the process of segmenting a given sentence into a row of morphemes. A morpheme is a minimal grammatical unit, such as a word or a suffix [20]. MeCab (Graduate School of Informatics Kyoto University, Kyoto, Japan, and NTT Communication Science Laboratories, Seika, Kyoto) performs morphological analysis using the hidden Markov model and statistical processing [21].

Morphological analysis of tweets and newspaper articles was performed for each week for 14 weeks (3 months and 1 week, from January 22 to April 29), covering the week before the publication of the STAP cell paper (Week 0) and the 13 weeks after (Weeks 1-13). The nouns identified in the tweets and newspaper articles through the analysis were extracted, and their frequencies were calculated.

The top 100 nouns that appeared most frequently each week in tweets and newspaper articles during the 14 weeks were then extracted. Terms that were either positive or negative toward STAP cells were then selected from among these 100 frequently appearing nouns. Changes in the use of these positive and negative terms each week were also investigated. Authors Yuya Sugawara, a specialist in medical informatics, and Hiroto Narimatsu, a medical doctor, selected the positive or negative Japanese terms semantically.

R version 2.13.0 (R Foundation for Statistical Computing) was used for morphological analysis. The R package used for our morphological analysis was RMeCab version 0.97. RMeCab is a package developed to operate the Japanese morphological analyzer MeCab version 0.98 from R.

Extraction of Accounts and Tweets

To investigate the attributes of Twitter account holders who tweeted using negative terms on STAP cells, 100 from the total 558 accounts that sent STAP-cells tweets from February 5 to 18, 2014, containing the negative term “unnatural” were randomly selected. Profile searches were performed on these 100 accounts using Twitter’s Profile Search feature [22]. All the 558 accounts were assigned random numbers. We extracted 100 accounts in the descending order of the assigned random numbers using Excel (Microsoft Corporation, Redmond).

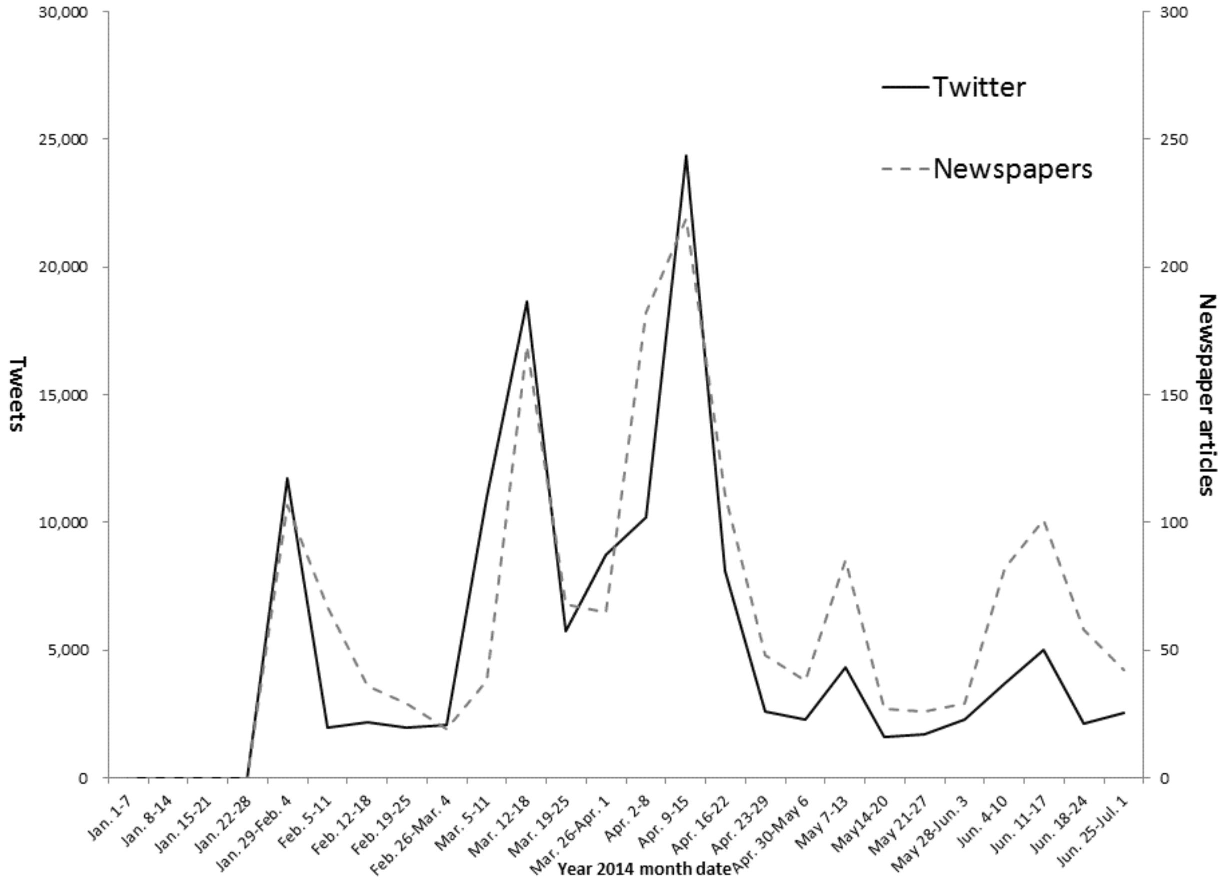
In order to classify the content of tweets on STAP cells, 1000 tweets were randomly extracted from the total 134,958 tweets that were assembled over the 6-month search period; all the 134,958 tweets were assigned random numbers, and the 1000 tweets were extracted in the descending order of the assigned numbers. These 1000 tweets were classified into four types: STAP related; links to news sites, and so on; false scientific or medical claims; and topics unrelated to STAP specified by Yuya Sugawara and Hiroto Narimatsu.

Results

Number of Newspaper Articles and Tweets Per Week

Figure 1 shows the numbers of newspaper articles and tweets containing the term “STAP cells” each week for the 6-month period following the publication of the STAP cells paper. A total of 1646 newspaper articles and 134,958 tweets appeared in 6 months. The number of both newspaper articles and tweets followed a similar trend, exhibiting transient increases during the periods from January 29 to February 4, from March 12 to 18, and from April 9 to 15. The numbers of tweets and newspaper articles were correspondingly 11,718 and 107 during the first period, 18,649 and 169 in the second period, and 24,344 and 219 in the third period.

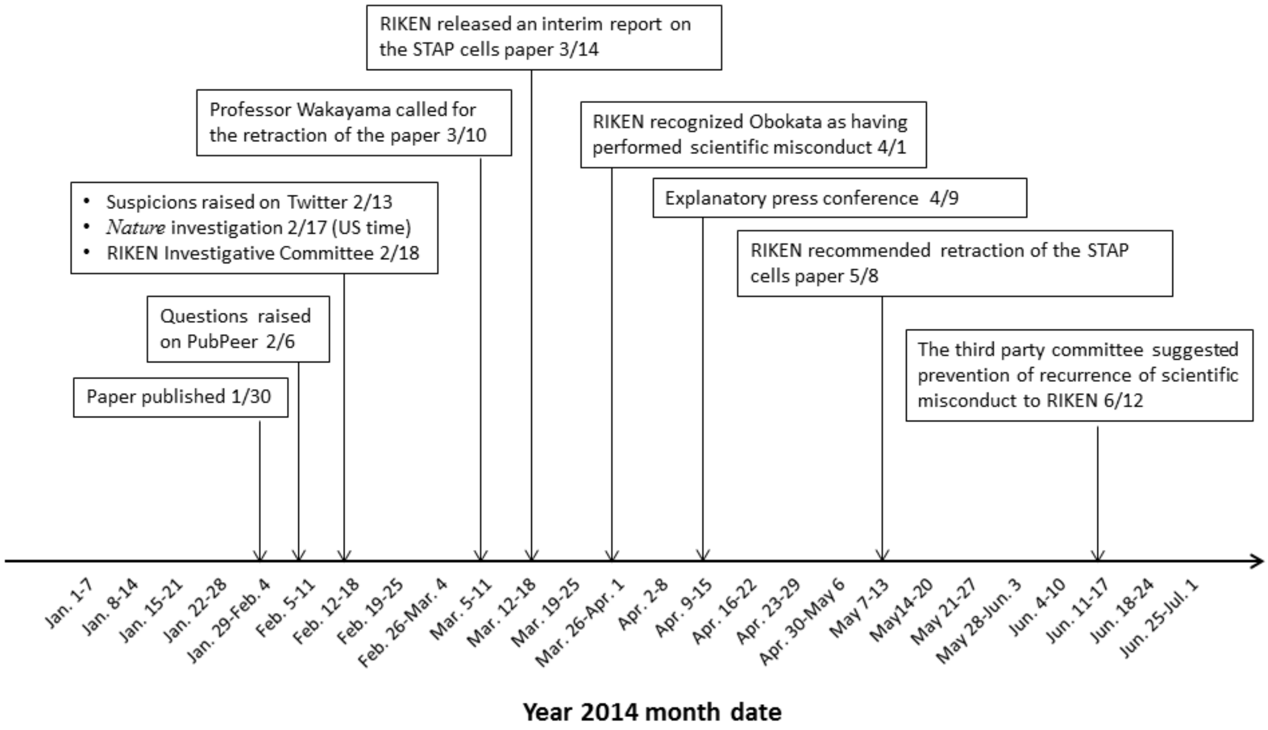
Figure 1. A comparison of the numbers of tweets and newspaper articles containing the word “STAP” between January 1 and July 1, 2014. The solid line shows the number of tweets and the broken line newspaper articles.



Timeline of the STAP Cells Scandal

Figure 2 shows the timeline of the STAP cells scandal. The paper on STAP cells was published on January 30, 2014 [23]. On March 14, RIKEN released an interim report on the paper [24]. Subsequently, on April 9, Haruko Obokata held a news conference [25].

Figure 2. Timeline of events related to the stimulus triggered acquisition of pluripotency (STAP) cells scandal.



Frequency of Nouns Appearing in Tweets and Newspaper Articles

We conducted a morphological analysis of the tweets and newspaper articles that appeared during the first 13 weeks (3 months) following the publication of STAP cells paper on January 30. The frequency of appearance of nouns in these tweets and newspaper articles during the same period was also investigated. A total of 100 nouns that appeared most frequently were extracted and classified whether positive or negative toward STAP cells. The positive terms selected were “major discovery” and “ground-breaking,” whereas the negative terms were “unnatural,” “fabrication,” and “falsification.”

Figures 3 and 4 show the frequency of use of positive and negative terms in tweets and newspaper articles during the first 3 months (13 weeks), including 1 week before paper publication (14 weeks in total), following the publication of the STAP cells study. In Twitter, the frequency of positive terms was the highest (n=432) in the period from January 29 to February 4, whereas the frequency of negative terms was the highest (n=835) in the period from February 12 to 18. The frequency of negative terms was 1.93 times higher than that of positive terms. In newspaper articles, the frequency of positive terms was the highest (n=31) from January 29 to February 4 and March 12 to 18, whereas the frequency of negative terms was at the maximum (n=296) from April 9 to 15. The frequency of negative terms increased 6 weeks after the STAP paper publication. The highest frequency of negative terms in newspaper articles appeared 8 weeks later compared with that for tweets.

Figure 3. Frequencies of the use of positive terms related to stimulus triggered acquisition of pluripotency (STAP) in Twitter and newspapers.

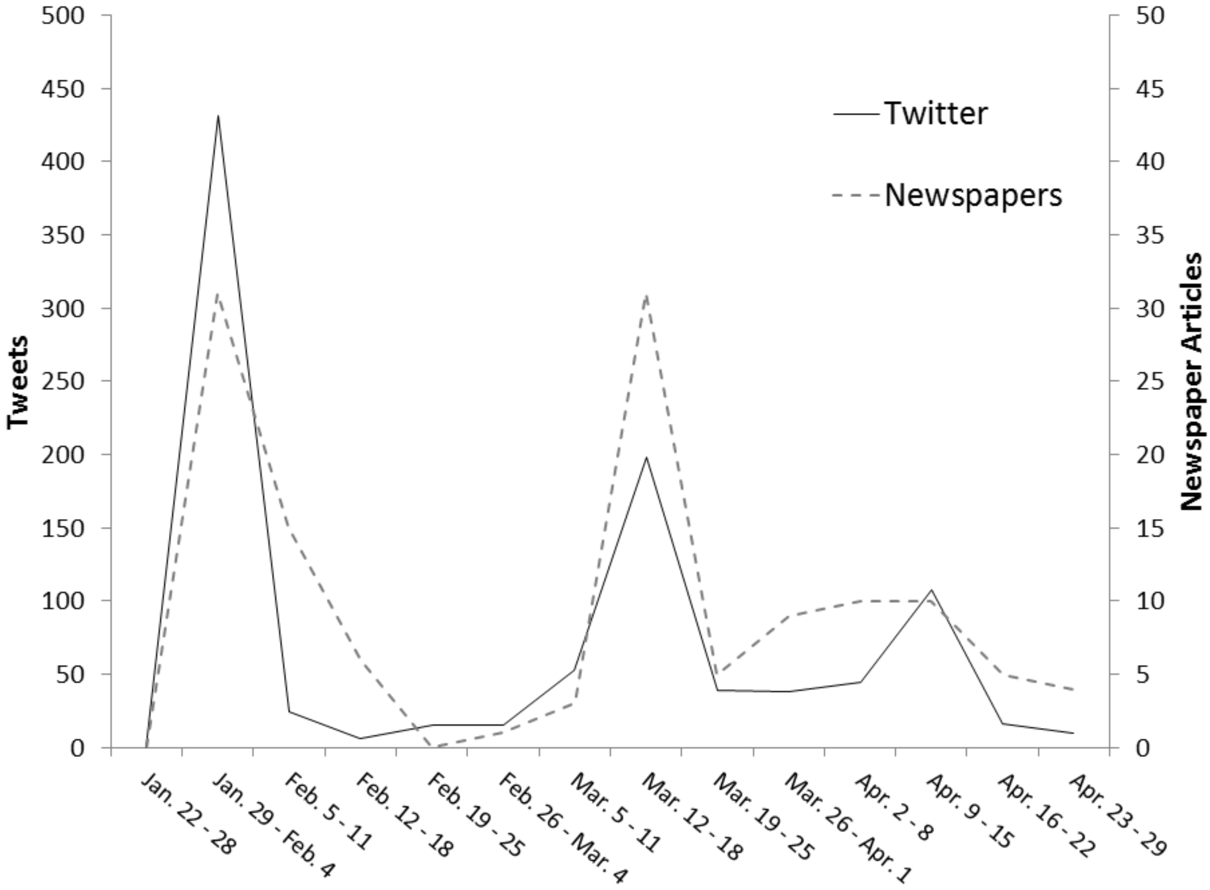
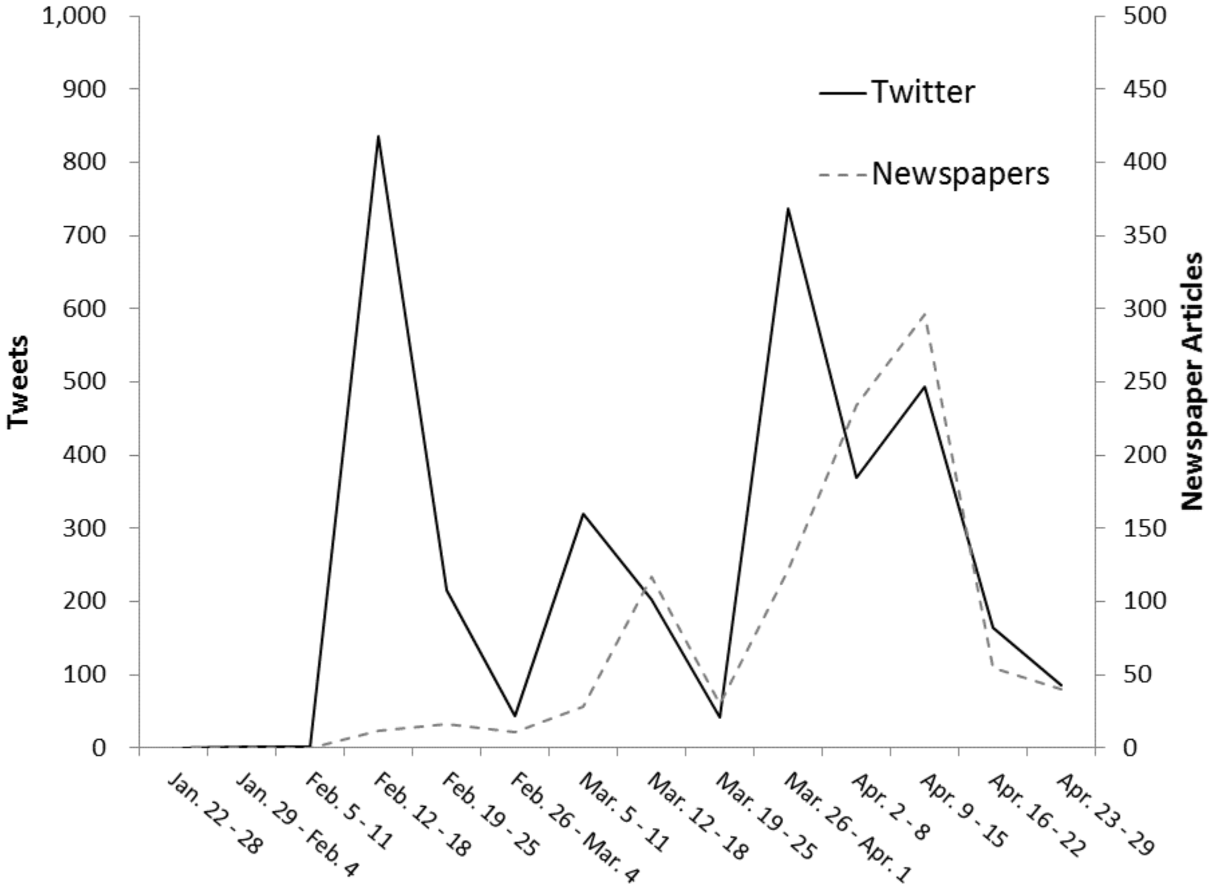


Figure 4. Frequencies of the use of negative terms related to stimulus triggered acquisition of pluripotency (STAP) in Twitter and newspapers.



Account Attributes and Tweet Content

Four of the account holders who published tweets on STAP cells containing the negative term “unnatural” were considered to possess certain specialist knowledge being a former biological researcher, a staff at a manufacturer of reagents for biological experiments, a science writer, and a science journalist. The others included 8 news sites, 9 bots (accounts that generate tweets automatically), and 79 accounts for which the level of specialist knowledge could not be determined.

Of the total 134,958 tweets generated during the 6-month period, 1000 tweets were randomly selected and found to contain STAP-related opinions (43.3%, 433/1000), links to news sites and other sources (41.4%, 414/1000), false scientific or medical claims (8.9%, 89/1000), and topics unrelated to STAP (6.4%, 64/1000). The examples of tweets that were evaluated to make false scientific or medical claims were as follows: “STAP cells will give us immortality,” “STAP cells are produced from the skin of newborn babies,” and “STAP cells can be made at home.”

Discussion

Principal Findings

This study revealed that the discussion on misconduct in the STAP cells affair was taken up at an earlier stage on Twitter than by newspapers. Positive terms appeared both on Twitter and in newspaper articles during the first week after STAP cells study was published, suggesting that the tendency was for STAP cells to be viewed positively. From Week 2 (February 5-11), however, the use of positive terms decreased and negative ones started to appear. On the Twitter, the frequency of negative terms increased during February 12-18 and starting again from March 19 to 25, 2 weeks after Obokata’s coauthor called for the retraction of the paper. In newspaper articles, the negative terms increased between March 19 and 25 and April 9 and 15; an explanatory press conference by Obokata was held during April 9 and 15. The highest use of negative terms in newspaper articles was observed during this week. The use of both positive and negative terms declined beginning during April 16-22. The story on suspicions were taken up by newspapers after the “unnatural” nature of STAP cells had been pointed out on Twitter and after RIKEN announced that it had found Obokata guilty of scientific misconduct, with the increased use of “fabrication” and “falsification” setting a different tone. This tone might have affected public opinion on the STAP cells paper. In both the Schön and Useok scandals, both of which constituted serious scientific misconduct, the retraction of papers took several years [26,27]. Obokata’s STAP cell paper, however, was withdrawn after only about 5 months [23,28]. New tools such as Twitter might have played a role in the early process leading to the retraction of the paper in the STAP cell scandal. Publishing on Twitter has the clear advantage that it is speedy [14]. The greatest advantage of using Twitter for scientific discussion is the rapid result of a debate. Scientific misconduct has continuously occurred after the Useok scandals. In Japan, Valsartan-related misconduct occurred [5,29-33], and after its publication in 2007 [34], a researcher expressed concerns about the study [35]. The misconduct by an employee of a pharmaceutical company was revealed and the paper was retracted in 2013. This misconduct occurred in the clinical trials, which would affect the treatment strategy in many patients. Thus, this misconduct is more serious than that in the STAP cells affair because the fallacious result would harm humans, whereas the STAP cells study was in a basic science. However, this misconduct seems to be perceived less seriously than the STAP cells affair.

Accuracy of Twitter is not always guaranteed. The medical and scientific accuracy of the tweets in this study was questionable for 8.9% of the cases. Accuracy is a constant issue not only in social media but also in Web-based information in general. Caution is always required when using Twitter and other Web-based sites to identify the wrong information. The Japan Internet Medical Association has issued a guide on using Web-based medical information [36]. This guide states that when using medical or health-related information taken from the Internet, members of the public should check that the source is clearly named, that it is backed up by objective evidence, and that the information has been provided by a public medical facility or research institution. If the provider of the information checks the originator or the identity of the account retweeting the information, it may be possible to evaluate its objectivity on the basis of factors such as links included in tweets. To a certain extent, judging the accuracy of information on Twitter is feasible.

The greatest advantage of using Twitter for scientific discussion is the rapidity with which the debate proceeds. However, Twitter also contains inaccurate information and excessive arguments that may become abusive. The latter can have a restraining effect on researchers. Depending on how the functions of Twitter are used, it may be possible to distinguish inaccurate information and to avoid abuse to some extent. Hence, the benefits of the rapid discussion enabled by Twitter, as shown in this study, can be enjoyed while limiting the risks of its disadvantages.

Limitations

This study showed the advantages of carrying out a scientific discussion on Twitter, but the scope of the study was limited. First, only top tweets were analyzed and an analysis of all tweets might have revealed different views. Second, the content of the discussion on Twitter, particularly on whether a discussion on misconduct in the STAP cells study took place, was not scrutinized. Third, other forms of media such as blogs, weekly magazines, and television were not investigated. Inaccurate images and articles broadcast by other media may also have had an effect on Twitter. Moreover, we did not analyze the accuracy of newspapers. The accuracy of information in newspapers was not necessarily better than that of Twitter. It is possible that the newspaper articles contained inaccurate statements.

The discussion of misconduct in the STAP cells study might have spread rapidly as it involved copying and pasting, a careless behavior familiar to the public. Specialized and complicated misconduct would be less likely discussed by the public on Twitter.

Conclusions

The discussion on scientific misconduct involving the STAP cell study took place at an earlier stage on Twitter than in newspapers, representatives of the traditional media. Results of the study suggest that the Twitter debate might have contributed to the resolution of the STAP cell scandal.

Conflicts of Interest

None declared.

References

1. Tijdink JK, Verbeke R, Smulders YM. Publication pressure and scientific misconduct in medical scientists. J Empir Res Hum Res Ethics 2014 Dec;9(5):64-71. [doi: 10.1177/1556264614552421] [Medline: 25747691]
2. Oksvold MP. Incidence of Data Duplications in a Randomly Selected Pool of Life Science Publications. Sci Eng Ethics 2016 Apr;22(2):487-496. [doi: 10.1007/s11948-015-9668-7] [Medline: 26065681]
3. Noyori R, Richmond JP. Ethical Conduct in Chemical Research and Publishing. Adv Synth Catal 2013 Jan 08;355(1):3-9. [doi: 10.1002/adsc.201201128]
4. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. Proc Natl Acad Sci U S A 2012 Oct 16;109(42):17028-17033 [FREE Full text] [doi: 10.1073/pnas.1212247109] [Medline: 23027971]
5. Tanimoto T. A perspective on the benefit-risk assessment for new and emerging pharmaceuticals in Japan. Drug Des Devel Ther 2015;9:1877-1888 [FREE Full text] [doi: 10.2147/DDDT.S62636] [Medline: 25848223]
6. The Huffington Post. Reform Committee indicates that the STAP cell paper, the world's worst misconduct, is to be included in textbooks URL: http://www.huffingtonpost.jp/2014/06/14/stap-cells-text_n_5494275.html [accessed 2016-09-24] [WebCite Cache ID 6kkWE2IBA]
7. Osterath B. Deutsche Welle. 2014. From stem cells to physics fraudulent science results are plenty but hard to find URL: http://www.dw.com/en/from-stem-cells-to-physics-fraudulent-science-results-are-plenty-but-hard-to-find/a-17687505 [accessed 2016-05-19] [WebCite Cache ID 6hcGeMbKn]

8. Brumfiel G. Misconduct finding at Bell Labs shakes physics community. Nature 2002 Oct 3;419(6906):419-421. [doi: 10.1038/419419a] [Medline: 12368818]
9. Murakami K. Effectiveness of and issues with electronic journals in light of the ES cell paper fabrication affair [In Japanese]. Current Awareness 2006;287:2-4 [FREE Full text]
10. Uchimura N. STAP Cell and Science Communication. JJSC 2014;15:161 [FREE Full text]
11. Normile D, Vogel G. Cell biology. STAP cells succumb to pressure. Science 2014 Jun 13;344(6189):1215-1216. [doi: 10.1126/science.344.6189.1215] [Medline: 24925996]
12. Mandavilli A. Peer review: Trial by Twitter. Nature 2011 Jan 20;469(7330):286-287. [doi: 10.1038/469286a] [Medline: 21248816]
13. Faulkes Z. The vacuum shouts back: postpublication peer review on social media. Neuron 2014 Apr 16;82(2):258-260 [FREE Full text] [doi: 10.1016/j.neuron.2014.03.032] [Medline: 24742455]
14. Jogalekar AS. Social Media, Peer Review, and Responsible Conduct of Research (RCR) in Chemistry: Trends, Pitfalls, and Promises. Account Res 2015 Jul;22(6):402-430 [FREE Full text] [doi: 10.1080/08989621.2015.1047705] [Medline: 26155733]
15. Yeo SK, Liang X, Brossard D, Rose KM, Korzekwa K, Scheufele DA, et al. The case of #arseniclife: Blogs and Twitter in informal peer review. Public Underst Sci 2016 May 26:1-16. [doi: 10.1177/0963662516649806] [Medline: 27229853]
16. Twitter. Twitter Help Center: FAQs about top search results URL: https://support.twitter.com/articles/131209 [accessed 2016-05-19] [WebCite Cache ID 6hcGr7s8S]
17. Yomiuri. Japan's Best-Read Newspaper Yomiuri Media Data 2016-2017 URL: http://adv.yomiuri.co.jp/m-data/english/download/ymd_2016-2017.pdf [accessed 2016-08-20] [WebCite Cache ID 6jtkLJ7nL]
18. Sankei. The Sankei Shimbun Media Data 2014 URL: http://www.sankei-ad-info.com/pdf/media_data2014_e.pdf [accessed 2016-08-14] [WebCite Cache ID 6jkTj5WwM]
19. Nikkei Telecom (in Japanese). URL: http://telecom.nikkei.co.jp/ [accessed 2016-05-19] [WebCite Cache ID 6hcH0UgHL]
20. Uchimoto K, Sekine S, Isahara H. The unknown word problem: a morphological analysis of Japanese using maximum entropy aided by a dictionary. Proceedings of EMNLP 2001:91-99.
21. Hiekata K, Yamato H, Tsujimoto S. Ontology based knowledge extraction for shipyard fabrication workshop reports. Expert Syst Appl 2010:7380-7386.
22. Twitter Profile Search (in Japanese). URL: http://tps.lefthandle.net/ [accessed 2016-05-19] [WebCite Cache ID 6hcHDKnt9]
23. Review: The STAP Issue (Volume 1) [In Japanese]. In: The Yomiuri Shimbun Morning edition. Tokyo: The Yomiuri Shimbun; Feb 15, 2015.
24. Riken. 2014. Interim report on the investigation of the Obokata et al. articles URL: http://www.riken.jp/en/pr/press/2014/20140314_1/ [accessed 2016-11-27] [WebCite Cache ID 6mK044vwp]
25. Otake T. The Japan Times. 2014. Obokata says STAP cell discovery not fabrication, claims Riken dissuaded her from giving her side of story earlier URL: http://www.japantimes.co.jp/news/2014/04/09/national/obokata-says-stap-cell-discovery-not-fabrication-claims-riken-dissuaded-her-from-giving-her-side-of-story-earlier/ [accessed 2017-02-16]
26. Bao Z, Batlogg B, Berg S, Dodabalapur A, Haddon RC, Hwang H, et al. Retraction. Science 2002 Nov 01;298(5595):961-961. [doi: 10.1126/science.298.5595.961b]
27. Sciencemag. Special Online Collection: Hwang et al. Controversy -- Committee Report, Response, and Background URL: http://www.sciencemag.org/site/feature/misc/webfeat/hwang2005/ [accessed 2016-09-24] [WebCite Cache ID 6kkWoCuOs]
28. Obokata H, Wakayama T, Sasai Y, Kojima K, Vacanti MP, Niwa H, et al. Stimulus-triggered fate conversion of somatic cells into pluripotency. Nature 2014 Jan 30;505(7485):641-647. [doi: 10.1038/nature12968] [Medline: 24476887]
29. McCurry J. Former Novartis employee arrested over valsartan data. The Lancet 2014 Jun;383(9935):2111. [doi: 10.1016/S0140-6736(14)61015-1]
30. Normile D. Tampered data cast shadow on drug trial. Science 2013 Jul 19;341(6143):223. [doi: 10.1126/science.341.6143.223] [Medline: 23868992]
31. Normile D. Biomedicine. Faulty drug trials tarnish Japan's clinical research. Science 2014 Jul 4;345(6192):17. [doi: 10.1126/science.345.6192.17] [Medline: 24994629]
32. Yui Y. Prevention of misconduct in clinical trials in Japan. Lancet 2014 Sep 20;384(9948):1098. [doi: 10.1016/S0140-6736(14)61675-5] [Medline: 25241720]
33. Tanimoto T, Kami M, Shibuya K. Research misconduct and scientific integrity: a call for a global forum. Lancet 2013 Sep 14;382(9896):940. [doi: 10.1016/S0140-6736(13)61933-9] [Medline: 24034298]
34. Mochizuki S, Dahlöf B, Shimizu M, Ikewaki K, Yoshikawa M, Taniguchi I, Jikei Heart Study group. Valsartan in a Japanese population with hypertension and other cardiovascular disease (Jikei Heart Study): a randomised, open-label, blinded endpoint morbidity-mortality study. Lancet 2007 Apr 28;369(9571):1431-1439. [doi: 10.1016/S0140-6736(07)60669-2] [Medline: 17467513]
35. Yui Y. Concerns about the Jikei Heart Study. Lancet 2012 Apr 14;379(9824):e48. [doi: 10.1016/S0140-6736(12)60599-6] [Medline: 22500880]

36. JIMA. Guide to the Use of Medical Information on the Internet [In Japanese] URL: http://www.jima.or.jp/userguide1.html [accessed 2016-05-19] [WebCite Cache ID 6hcHKpsDG]

Abbreviations

STAP: Stimulus-Triggered Acquisition of Pluripotency

Edited by G Eysenbach; submitted 13.10.16; peer-reviewed by K Bazdaric, Z Faulkes; comments to author 03.11.16; revised version received 08.12.16; accepted 08.02.17; published 28.02.17.

Please cite as:
Sugawara Y, Tanimoto T, Miyagawa S, Murakami M, Tsuya A, Tanaka A, Kami M, Narimatsu H
Scientific Misconduct and Social Media: Role of Twitter in the Stimulus Triggered Acquisition of Pluripotency Cells Scandal
J Med Internet Res 2017;19(2):e57
URL: http://www.jmir.org/2017/2/e57/
doi:10.2196/jmir.6706
PMID:28246071

Original Paper

Perceptions of Menthol Cigarettes Among Twitter Users: Content and Sentiment Analysis

Shyanika W Rose[1], MA, PhD; Catherine L Jo[2], MSPH, PhD; Steven Binns[3], MPH; Melissa Buenger[3], MPH; Sherry Emery[3], PhD; Kurt M Ribisl[2,4], PhD

[1]The Schroeder Institute for Tobacco Research and Policy Studies at Truth Initiative, Washington, DC, United States
[2]Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[3]Health Media Collaboratory, NORC at the University of Chicago, Chicago, IL, United States
[4]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Shyanika W Rose, MA, PhD
The Schroeder Institute for Tobacco Research and Policy Studies at Truth Initiative
900 G Street, NW
4th Floor
Washington, DC, 20001
United States
Phone: 1 202 454 5917
Fax: 1 202 454 5555
Email: srose@truthinitiative.org

Abstract

Background: Menthol cigarettes are used disproportionately by African American, female, and adolescent smokers. Twitter is also used disproportionately by minority and younger populations, providing a unique window into conversations reflecting social norms, behavioral intentions, and sentiment toward menthol cigarettes.

Objective: Our purpose was to identify the content and frequency of conversations about menthol cigarettes, including themes, populations, user smoking status, other tobacco or substances, tweet characteristics, and sentiment. We also examined differences in menthol cigarette sentiment by prevalent categories, which allowed us to assess potential perceptions, misperceptions, and social norms about menthol cigarettes on Twitter. This approach can inform communication about these products, particularly to subgroups who are at risk for menthol cigarette use.

Methods: Through a combination of human and machine classification, we identified 94,627 menthol cigarette-relevant tweets from February 1, 2012 to January 31, 2013 (1 year) from over 47 million tobacco-related messages gathered prospectively from the Twitter Firehose of all public tweets and metadata. Then, 4 human coders evaluated a random sample of 7000 tweets for categories, including sentiment toward menthol cigarettes.

Results: We found that 47.98% (3194/6657) of tweets expressed positive sentiment, while 40.26% (2680/6657) were negative toward menthol cigarettes. The majority of tweets by likely smokers (2653/4038, 65.70%) expressed positive sentiment, while 91.2% (320/351) of nonsmokers and 71.7% (91/127) of former smokers indicated negative views. Positive views toward menthol cigarettes were predominant in tweets that discussed addiction or craving, marijuana, smoking, taste or sensation, song lyrics, and tobacco industry or marketing or tweets that were commercial in nature. Negative views toward menthol were more common in tweets about smoking cessation, health, African Americans, women, and children and adolescents—largely due to expression of negative stereotypes associated with these groups' use of menthol cigarettes.

Conclusions: Examinations of public opinions toward menthol cigarettes through social media can help to inform the framing of public communication about menthol cigarettes, particularly in light of potential regulation by the European Union, US Food and Drug Administration, other jurisdictions, and localities.

(J Med Internet Res 2017;19(2):e56) doi:10.2196/jmir.5694

KEYWORDS

tobacco products; menthol; smoking; social media; Twitter messaging; policy; public opinion

Introduction

Menthol is a characterizing flavor that imparts a minty flavor and a cool sensation masking the harshness of cigarette smoke [1,2]. Menthol cigarettes comprised 31% of market share for cigarettes sold in the United States in 2012 [3]. Use of menthol versus nonmenthol cigarettes is higher among African American, female, younger, and lesbian, gay, bisexual, or transgender (LGBT) smokers [4-8]. For instance, among smokers, 84% of African Americans versus 24% of whites smoke menthol cigarettes [6], and adolescent (57%) and young adults (45%) use menthol cigarettes at higher rates than all ages of older smokers [8]. All of these groups have been targeted by tobacco industry menthol cigarette marketing [9,10].

The US Food and Drug Administration (FDA) banned characterizing flavors in cigarettes except for menthol or tobacco flavor in 2009 [11]. Menthol was exempted pending further review. In 2011, the FDA's Tobacco Products Science Advisory Committee (TPSAC) issued a report concluding that "Removal of menthol cigarettes would benefit public health in the United States" [12]. TPSAC and a subsequent FDA report [13] concluded that menthol cigarette use is likely associated with increased smoking initiation, increased levels of nicotine addiction, and reduced smoking cessation success, particularly among African Americans. Moreover, menthol cigarettes present the same disease risk as nonmenthol cigarettes [14], with smoking causing up to two-thirds of deaths in smokers [15]. The FDA has not yet decided on whether to ban menthol cigarettes [16]. However, other jurisdictions are moving forward with menthol cigarette bans and restrictions. Brazil became the first country to pass legislation banning menthol and flavors from tobacco products in 2012, but the ban has not yet taken effect [17]. In 2014, the city of Chicago enacted a ban on menthol and flavored tobacco product sales near schools. The European Union (EU) also banned menthol cigarettes with a 2020 implementation date [17]. In 2015, some Canadian provinces prohibited flavored and menthol cigarettes and tobacco sales [17].

Understanding perceptions of menthol cigarettes given this changing global regulatory environment is critical to understanding how to best communicate to the general public about these products prior to widespread policy implementation. Perceptions of menthol cigarettes among both smokers and the general public may contribute to use [9,18-21]. A study of US adolescents and young adults found that endorsing promenthol cigarette beliefs (positive sentiment) was associated with intention to use menthol cigarettes among noncurrent menthol cigarette users and intention to use other tobacco products among current menthol cigarette smokers not currently using other products [22].

Commonly reported positive perceptions of menthol cigarettes are that they are more cooling or refreshing than nonmenthol cigarettes [18,22], have a medicinal effect when the smoker is sick [18,23], and are part of African American culture or tradition [24]. Beliefs that menthol cigarettes are less harmful than nonmenthol cigarettes appear more common among African American smokers [18]. In negative sentiments, among current smokers, 22% believe that menthol cigarettes are more addictive, while 12.1% believe that menthol cigarettes are harder to quit smoking than nonmenthol cigarettes [21].

Most studies of perceptions of menthol cigarettes have been conducted through surveys or focus groups [18,21,23-26]. Analyzing the amount and content of social media messages is a relatively new and promising way to gather information about health-related topics [27,28]. On Twitter, users can share 140-character messages called *tweets* with their followers and the general public. The forum is popular, with the number of daily tweets topping 500 million [29]. An estimated 23% of the US adult online population used Twitter in 2014 [30]. Twitter also reaches a global audience with an estimated 316 million active users worldwide [31].

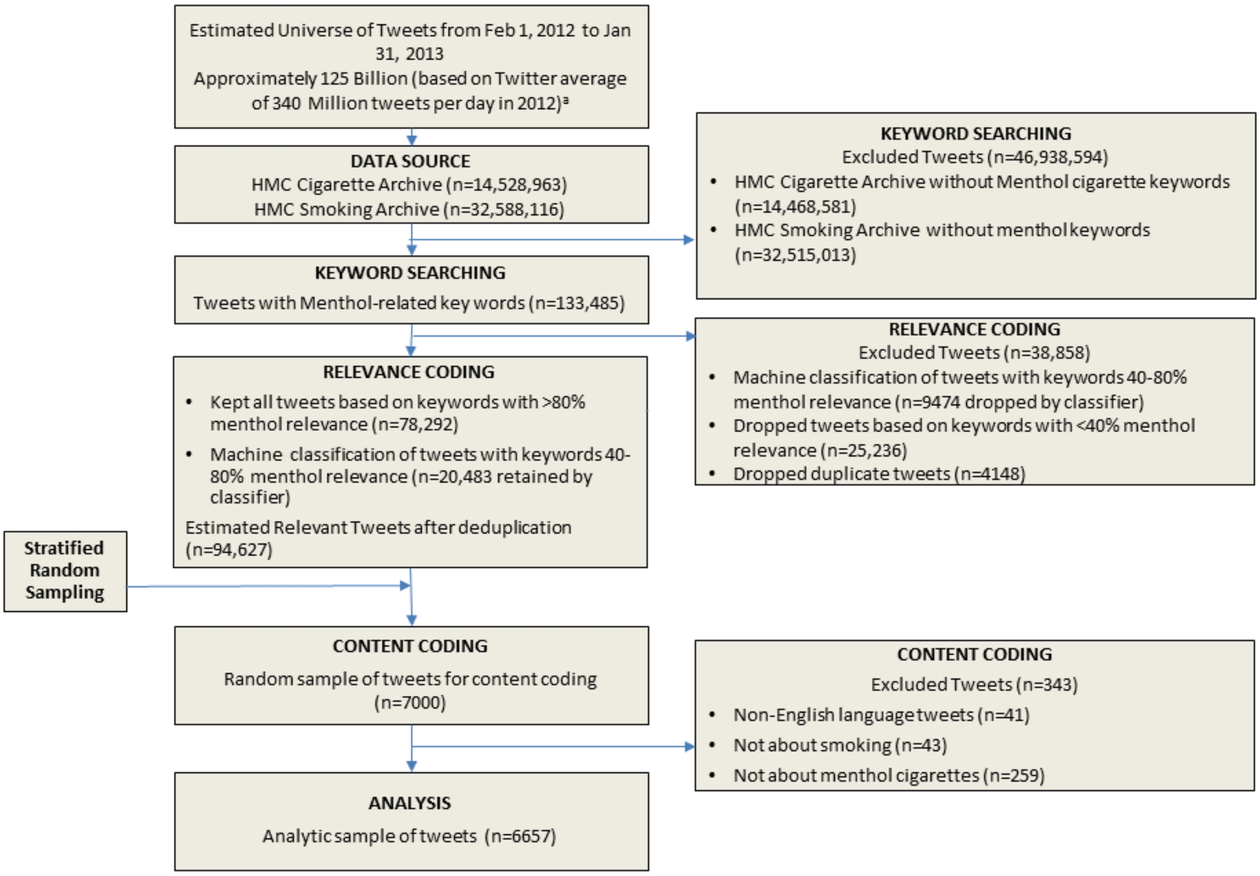
Used disproportionately by younger and racial and ethnic minority populations who also disproportionately use menthol cigarettes [30], Twitter may provide insight into use and perceptions of these products. Several studies have examined Twitter conversations about electronic cigarettes [32-36] and hookah [34,37]. However, to our knowledge, no study has examined Twitter conversations about menthol cigarettes. As in prior Twitter studies of tobacco, we used a content analysis approach [35-37] and coded specifically for sentiment [35,38] toward menthol cigarettes. Our purpose was to identify the content and frequency of conversations about menthol cigarettes, including themes, populations, user smoking status, other tobacco or substances, tweet characteristics (eg, song lyrics or pop culture references, commercial tweets), and sentiment. We also examined differences in menthol cigarette sentiment by prevalent categories, which allowed us to assess perceptions and potential misperceptions and social norms about menthol cigarettes on Twitter. This approach can inform communication about these products, particularly to subgroups who are at risk for menthol cigarette use.

Methods

Design

We conducted a search to define a longitudinal set of tweets related to menthol cigarettes between February 1, 2012 and January 31, 2013 (shown in Figure 1 [39]). This time period provides a year of data and covers a time period prior to regulatory action on menthol cigarettes globally.

Figure 1. Overall study design and process for collecting menthol-relevant tweets for coding and analysis.^{a}Data source: Internet Live Stats. HMC: Health Media Collaboratory.



Data Source

This observational study used data collected by the Health Media Collaboratory (HMC) [40]. The HMC accessed Twitter status updates through the Gnip PowerTrack Firehose (Gnip, Inc) [41], which prospectively collects all Tweets from public Twitter feeds.

The HMC created two archives that collected all Twitter status updates containing the keywords “cig” or “cigarette” (known as the cigarette archive) or the keyword “smoking” (the smoking archive). We chose those keywords after an extensive process of precision testing of terms likely to obtain tobacco-related content (Health Media Collaboratory, *Procedures Manual*, unpublished, 2012). The Gnip PowerTrack collected the content of the tweet, metadata such as username (who made the tweet), date and time the update was made, and hashtags (user-generated descriptive tags).

Keyword Searching

At the time of our search for menthol cigarettes, the smoking archive had 32,588,116 tweets and the cigarette archive had 14,528,963. We identified an initial set of keywords including general terms, major menthol cigarette brands, slang terms, and common misspellings that we expected to generate relevant menthol cigarette-relevant content (see Table 1). We filtered the smoking and cigarette archives for each term, resulting in 133,485 tweets.

Table 1. Keywords used to identify menthol cigarette-related tweets and assessment of relevance.

Keywords		Cigarette archive		Smoking archive	
		Number of tweets found	% of relevant tweets in 100-tweet sample or all if <100 with keyword	Number of tweets found	% of relevant tweets in 100-tweet sample or all if <100 with keyword
General terms					
	Menthol	22435[a]	52[a]	6897	93
	Mentholated	161	96	47	96 (n=45)
	Menthols	1388	100	2956	100
Brands					
	Camel Crush	1254	100	590	100
	Kool	3061[b]	35[b]	7966[b]	17[b]
	Kools	345	100	1101	100
	Newport	12553	88	22645	98
	Newports	5985	100	15709	98
	Salem	4078[b]	20[b]	803[b]	12[b]
	Salems	33[b]	100 (n=33)	94	91 (n=86)
	Skyline (Marlboro)	174[b]	28[b]	617[b]	5[b]
Misspelling					
	Methol	98	96 (n=94)	61	95 (n=58)
	Methols	23	96 (n=22)	26	100 (n=26)
Slang					
	Newp	125	100	322	98
	Newps	132	100	307	100
	Port	6972[b]	1[b]	7522[a]	77[a]
	Ports	1565[b]	26[b]	5440	100

[a] Tweets were subject to machine learning assessment.

[b] Tweets were dropped as having <40% relevant content. All other tweets were retained, as keyword precision tests were ≥80% menthol cigarette related based on relevance coding.

Relevance Coding

Once the searches were completed, 2 coders independently coded 100 randomly selected tweets from each keyword or all tweets if less than 100 for menthol relevance. We calculated a kappa statistic for each term to assess interrater reliability. For any term with at least .8 kappa score, one author (SWR) verified the coding and generated a measure of precision (percentage of identified tweets relevant to menthol cigarettes). For any keyword less than .8 kappa score, we coded another 100 tweets until we reached this level of interrater reliability. Of the 42 keywords searched, we achieved a kappa of at least .8 on the first attempt in 41 (98%) keywords.

We used DiscoverText (Texifter, LLC), a cloud-based text analytic software with active machine learning algorithms, allowing human coders to “train” the computer on a coding schema using a selection of tweets. If the keyword generated less than 40% of relevant tweets from the 100-tweet sample, we dropped all tweets with that keyword. If a keyword generated 80% or more relevant tweets in the 100-tweet sample, we kept all tweets. If the keyword generated 40% to 80% relevant tweets, we used the machine learning to assign a probability of a given tweet being menthol cigarette relevant, validated the machine estimates using a second 100-tweet sample coded by human coders, and retained only tweets with greater than 0.8 probability of being menthol cigarette relevant based on the machine classification. Of the 29,957 tweets assessed by the classifier, we retained 68.37% (n=20,483).

After merging and deduplicating these sets of tweets across archives and keywords, we ended with 94,627 unique tweets to include in content coding.

Content Coding

We developed a codebook (see Table 2 for brief description of codes; Multimedia Appendix 1 provides the broader codebook) for directed content analysis [42], building on empirically derived constructs but allowing for emergent categories in the following areas: themes, populations, other tobacco or substances, tweet characteristics, user smoking status, and sentiment toward menthol cigarettes.

Table 2. Brief description of codes.

Categories	Definitions
Themes (topical content dealing with perceptions of menthol cigarettes)	
Taste/sensation	Taste, smell, or sensation (eg, cooling, minty, refreshing) of menthol cigarettes.
Health	Perceived health benefits or harms of menthol cigarettes, or beliefs about medicinal effects of menthol cigarette use.
Cessation	Desire or lack of desire to quit smoking menthol cigarettes; beliefs about whether menthol cigarettes are harder or easier to quit than nonmenthol cigarettes.
Addiction	Addiction, cravings, or desire for their lack of menthol cigarettes, beliefs that menthol cigarettes are more or less addictive than nonmenthol cigarettes.
Smoking behavior	Act or process of smoking menthol cigarettes, including time or place of smoking.
Tobacco control policies	Tobacco control policies or the impact of tobacco control policies.
Industry/marketing	Advertising, promotion, labeling, or packaging of menthol cigarettes, including references to tobacco companies and protobacco organizations.
User smoking status (likely smoking status of the Twitter user)	
Smoker	User writing tweet is likely to be a smoker.
Former smoker	User writing tweet is likely to be a former smoker.
Nonsmoker	User writing tweet is likely to be a nonsmoker. Also use this code for tweets from antitobacco organizations.
Unknown	Cannot determine smoking status of user writing tweet.
Tweet characteristics (additional information about tweets)	
Commercial	Branded promotional messages; URLs linking to commercial websites; usernames indicating affiliations with commercial sites; or the user's Twitter page consisting only of promotional tweets.
Song lyrics/pop culture	Tweets that are lyrics from songs about menthol cigarettes or pop culture references. This could be hashtags with an artist's name or popular television shows.
Other tobacco or substances (tweets mentioning use of other tobacco or other substances)	
Marijuana	Mentions of marijuana (eg, loud, blunt, weed, spliff, mary jane, wax).
Alcohol	Mentions of alcohol.
Cigars/little cigars and cigarillos	Mentions of cigars, little cigars or cigarillos, including blunts or specific brands.
Hookah	Mentions of hookah, waterpipe, shisha, or narghile for smoking tobacco or brands.
Nicotine replacement therapy	Mentions of nicotine replacement therapy such as nicotine gum, patch, or lozenge, or specific brands.
E-cigarette	Mentions of e-cigarettes, vaporizers, e-hookah, vape pens, etc, or of specific brands.
Smokeless/snus	Mentions of smokeless tobacco or snus (eg, dip, chew, snuff, spit) or specific brands.
Other tobacco or substances	Write-in mentions of other tobacco products (eg, pipe, roll-your-own) or brands, or other drugs (eg, LSD[a], cocaine).
Populations (groups associated with menthol cigarette use and menthol cigarette marketing)	
African Americans	African Americans or African American culture, image, or tradition or references to specific African American individuals in relation to menthol cigarettes.
LGBT[b]	LGBT populations or LGBT culture, image, or tradition or references to specific LGBT individuals in relation to menthol cigarettes.
Women	Women and menthol cigarettes or references to specific women and menthol cigarettes.
Children/adolescents	Children and adolescents, minors, or underage smokers (<18 years in the United States) and menthol cigarettes.
Other population	Write-in references to populations of people, does not include job categories (eg, rappers).
Sentiment (attitudes toward menthol cigarettes)	

Categories		Definitions
	Positive	Positive sentiment about menthol cigarettes.
	Negative	Negative sentiment about menthol cigarettes.
	Neutral	Either no sentiment about menthol cigarette or mixed sentiments (both positive and negative) about menthol cigarettes.

[a] LSD: lysergic acid diethylamide.

[b] LGBT: lesbian, gay, bisexual, or transgender.

We used the codebook to code a random sample of 7000 tweets weighted by prevalence of keywords in the menthol cigarette-relevant dataset. Thus, even tweets with low-prevalence keywords had a probability of being selected equal to their prevalence in the larger dataset. Retweets were coded the same as the original tweet. We coded the entire tweet content of modified retweets, but coded sentiment on the modification alone. Each tweet was coded with as many categories as were applicable.

Next, 4 coders independently coded 500 tweets, reconciling after each 100 and clarifying the codebook, as needed, including adding or broadening codes. Coders also used “other” (write-in) responses to code for new categories that were not in the initial codebook. We added prevalent other categories to the codebook, and reviewed previously coded tweets for the presence of any applicable other categories. Once coders reached kappa of .8 reliability across raters and categories, individual coders coded the remaining tweets independently. Coders flagged unclear tweets, which were adjudicated by a second coder (SWR, CLJ, or MB). We excluded non-English language tweets and those that contained relevant keywords but were not about menthol cigarettes (eg, smoking on Newport Beach, or enjoying Newport, RI).

Themes

We identified potential themes and the topical content of the tweet related to perceptions of menthol cigarettes, based on prior research [18-21,43], and refined these themes during preliminary coding. For example, after preliminary coding, we added categories such as smoking behavior and tobacco industry marketing. We also broadened some categories that were not easily distinguishable in the open-ended tweet format compared with a closed-ended survey, for instance, collapsing medicinal effect and harm perceptions into a broader health category. Additionally, we derived one potential theme, African American image [18]) from research on a single subpopulation. We lacked information on user race/ethnicity, so we could not readily code information on users’ perception of menthol cigarettes and their cultural identity. Instead, to capture this concept, we coded references to several populations that may be associated with menthol cigarette use in a set of population codes (described below). These codes included the concept of image. Our final list of categories comprised taste or sensation, health, smoking behavior, cessation, addiction, tobacco control, tobacco industry and marketing, and other (low-prevalence categories written in by coders).

Populations

We coded menthol-related tweets if they specifically mentioned menthol cigarettes in relation to populations targeted by menthol marketing [9] and with a higher prevalence of menthol cigarette use [4,5]. Such populations included African Americans, LGBT populations, women, and children and adolescents. We also included an “other” category to capture other populations mentioned (eg, Latinos). Other groups were at low prevalence in the sample.

Other Tobacco or Substances

We coded tweets based on co-mentions of menthol cigarettes and alcohol, marijuana, other substances, or other tobacco products, including cigars, little cigars, cigarillos (including blunts), e-cigarettes, smokeless tobacco, and hookah. Perceptions of other tobacco and substances may interact with perceptions of menthol cigarettes and be associated with increased risk behaviors [44,45].

Tweet Characteristics

We coded 2 additional characteristics of tweets, including commercial versus noncommercial tweets based on a definition in a prior study (eg, branded promotional messages) [32]. Finally, we coded for whether the tweet included a song lyric or popular culture quotation (eg, a television show reference), a category that emerged from the data. We created lists of song lyrics with menthol cigarette references (eg, “Smoking mad Newports ’cause I’m due in court / For an assault that I caught in Bridgeport, New York”, from Everyday Struggle by The Notorious B.I.G.) and added them to the list as we identified new lyrics. After compiling a final list, we reviewed all coded tweets for known song lyrics.

User Smoking Status

We coded tweets for likely user smoking status (nonsmoker, former smoker, smoker, or unknown) because differences in perceptions of menthol cigarettes have been associated with smoking status in prior research [21]. Smokers were characterized by tweets discussing current cigarette smoking. Former smokers were characterized by tweets about past smoking or having quit smoking. Nonsmokers noted that they did not smoke or were opposed to smoking; antitobacco organization tweets were included. Users not mentioning their own smoking status in the tweet were coded as unknown.

Sentiment

We coded each tweet for sentiment toward menthol cigarettes. Sentiment categories were positive, negative, and neutral/unclear. If a tweet included both positive and negative sentiments (eg, wanting to quit menthols, but liking the taste), it was coded as neutral/unclear. Tweets mentioning actual smoking, or wanting or craving a menthol cigarette, without further comment were coded as positive toward menthol cigarettes.

Analysis

Of the 7000 coded tweets, we excluded 343 (4.90%) as being either non-English or irrelevant to cigarette smoking or menthol cigarettes, with 95.1% retrieval precision and 90.4% retrieval recall within the HMC archive [46], leaving an analytic sample of 6657 tweets.

For these tweets, we provide univariate statistics of the frequency of tweets and provide qualitative narrative about frequent or notable content in each category. When examining differences in sentiment by other categories, we conducted chi-square tests to assess statistically significant differences in positive or negative sentiment toward menthol cigarettes.

Results

Frequency Analysis

Table 3 shows the frequency of themes, populations, other tobacco or substances, tweet characteristics, user smoking status, and sentiment.

Table 3. Frequency of menthol cigarette-relevant tweets by category (n=6657).

Category		n	%
Themes			
	Smoking	3983	59.83
	Taste/sensation	1676	25.18
	Addiction	1097	16.48
	Cessation	584	8.77
	Health	571	8.58
	Tobacco industry/marketing	390	5.86
	Other theme	299	4.49
	Tobacco control	295	4.43
Population			
	African American	745	11.19
	Women	587	8.82
	Child/adolescents	131	1.97
	Other population	28	0.42
	Lesbian, gay, bisexual, or transgender	22	0.33
Other tobacco or substances			
	Marijuana	1104	16.58
	Alcohol	221	3.32
	Cigar/little cigars and cigarillos	218	3.27
	Other substance	59	0.89
	E-cigarette	38	0.57
	Hookah	10	0.15
	Smokeless	3	0.05
Tweet characteristics			
	Song lyric/pop culture	1118	16.79
	Commercial	91	1.37
User smoking status			
	Smoker	4038	60.66
	Unknown	2141	32.16
	Nonsmoker	351	5.27
	Former smoker	127	1.91
Sentiment			
	Positive	3194	47.98
	Negative	2680	40.26
	Neutral/unclear	783	11.76

Themes

The majority of tweets about menthol cigarettes referenced smoking, taste, or sensation (eg, smell, coolness). About 17% referenced addiction (1097/6657), including explicitly discussing addiction, craving, or addiction-related behaviors such as chain smoking. All other themes—including cessation, health, tobacco industry and marketing, and tobacco control—were represented in fewer than 10% of the tweets coded. Tobacco control tweets (295/6657, 4.43%) included discussion of media campaigns or policy options including tax or price, minors' access, smoke-free air laws, sales or marketing restrictions, or warning labels, or were tweets by tobacco control organizations. Fewer than 0.60% of tweets (40/6657) referenced a menthol cigarette ban; most were about potential EU action. Only 0.20% (13/6657) were by tobacco control organizations.

Populations

About 11% of tweets (745/6657) linked African Americans and menthol cigarette use. For African Americans, use of menthol cigarettes, especially Newport cigarettes, was viewed as linked with African American culture. This was particularly true for African American males. This linkage was viewed both positively (“@___: If you black and smoke cigarettes at least smoke Newports”) and negatively (“Why do black people have a fetish with Newport cigarettes?”). Stereotypical jokes or hashtags were also associated with this category (“#YouKnowYouBlack,” “#WaysToPissOffaBlackPerson”). Women were linked with menthol cigarettes in about 9% of tweets (587/6657). Linkage with menthol cigarettes was associated with negative attitudes toward women, such as the unattractiveness of women who smoke (“Hate when I see pretty girls smoking Newports, it really just ruins it”) or lack of class (“#ShesARatchetIf she smokes newports”). Few tweets referenced children and adolescents or LGBT populations.

Other Tobacco or Substances

Over 16% of tweets (1104/6657) referenced marijuana and menthol cigarettes. These tweets referenced co-use of marijuana and menthol, such as smoking a menthol cigarette to enhance the high of smoking marijuana, comparing excessive use of marijuana as similar to chain smoking menthol cigarettes, or assessing the relative harms or appeal of marijuana and menthol cigarettes. These tweets usually judged marijuana as less harmful and more appealing than menthol cigarettes. Fewer tweets (218/6657, 3.27%) focused on dual use of menthol cigarettes and cigar products such as Black and Mild or Swisher Sweets, or blunts used for smoking marijuana. Alcohol and menthol cigarettes were mentioned in 3.32% (221/6657) of the tweets and discussed smoking behavior and drinking. Other tobacco products were rarely mentioned in conjunction with menthol cigarettes.

Tweet Characteristics

Only 1.37% of tweets (91/6657) in the sample were commercial tweets, such as tweets offering links to coupons or discounts. Just under 17% of tweets (1118/6657) referencing menthol cigarettes were song lyrics or popular culture quotations (such as a game show sketch from *Chappelle’s Show*, “I Know Black People,” famously asking the question “Why do black people like menthol cigarettes?”). We identified 30 songs or quotations tweeted in the sample.

User Smoking Status

The majority of tweets were likely by smokers (4038/6657, 60.66%) or users with unknown status (2141/6657, 32.16%).

Sentiment

Sentiment toward menthol cigarettes in the tweets was mixed. Overall, 47.98% (3194/6657) expressed positive sentiment toward menthol cigarettes and 40.26% (2680/6657) of tweets expressed negative sentiment. The remainder were neutral or unclear.

Sentiment Analysis

Sentiment varied by category across themes, populations, other tobacco and substances, and tweet characteristics (omitting categories with <1% prevalence) Figure 2 shows. Based on a chi-square test, tweets coded as referencing themes of addiction (including concepts such as craving or desire for a menthol cigarette), smoking, taste or sensation, and industry or marketing; marijuana as an other substance; and tweet characteristics of commercial tweets and song lyrics were significantly more likely to be positive than negative toward menthol cigarettes. However, both smoking and taste or sensation themes had close to 40% negative sentiment. Positive smoking and taste sentiment included positive attributes of smoking (relaxing, relieving stress, taking smoke breaks) and enjoyment of menthol (pleasant minty taste, soothing or cooling sensation). Negative sentiment included negative mood (stress, anxiety, “bad nerves”), negative feelings about other people smoking (“#ItsNotCuteWhen you smoking ports <<<<”), or dislike of the taste or smell of menthol cigarettes (“Smoking a newport. #gross,” “I hate the scent of newports!! Cigarettes in general Smell terrible, but especially newports!”).

Tweets referencing themes of cessation and health, and populations of women, children and adolescents, and African Americans were significantly more likely to be negative toward menthol cigarettes. Almost no cessation-related tweets noted proven quit methods such as the use of FDA-approved cessation aids, talking with a physician, or using a quitline [47]. Many noted plans to quit after smoking a last cigarette or pack, or asked for support for their quit attempt (“Today is my last day smoking cigarettes....waving good bye Newports! Wish me luck!!”). A few noted trying to quit by using an e-cigarette (14/6657 tweets). A small number referenced making a quit attempt in relation to the US Centers for Disease Control and Prevention’s Tips from Former Smokers campaign (“Seriously those commercials of the people with the hole in their throat made me quit smoking I haven’t had a newport in days”) [48]. Some also tried to quit by switching to a nonmenthol cigarette (“I need to get off menthols. I need to quit but if I start smoking something disgusting it may be easier.”).

Most health tweets were negative, noting negative health consequences of smoking. However, positive tweets related to health expressed a misperception of menthol cigarettes as having medicinal use when sick (“Menthol cigarettes are my saviour with this cold”). We rarely found harm perceptions of menthol as less harmful than other cigarettes (“Oh natural menthol healthy cigarettes”), but the converse message that menthol was harmful was more prominent (“SMOKERS (menthol or nonmenthol) really WILL HAVE stroke. So better stop smoking.” “Menthol cigarettes? So they soothe your throat while giving you cancer? Kinda like being killed by a clown.”). Regarding populations, negative sentiment toward menthol cigarettes was driven by negative views of smoking among children and adolescents, women, and African American populations. Tweets referencing cigars, alcohol, and tobacco control had no significant differences between positive and negative sentiment toward menthol cigarettes.

Among smokers, 65.70% (2653/4038) expressed positive sentiment toward menthol (“Listening to music and smoking a port so relaxed”) (Figure 3). Only 27.34% of smokers (1104/4038) reported negative sentiment toward cigarettes in general (“I cant be doin these menthols man.! Gotta Quit Smoking period. Maybe tomorrow”) or specifically toward menthol-type cigarettes or a menthol cigarette brand (“Had a mayfair menthol, its like smoking mouthwash”). Smoker conversations about menthol cigarettes were dominated by discussion of smoking behavior, addiction or craving, taste preference, and marijuana.

Nonsmokers (320/351, 91.2%) held predominately negative sentiments about menthol cigarettes and smoking in general (“All the girls at my job be smoking them Newports and had the nerve to offer me one. No ma’am, never.”). Former smokers were generally negative about menthol (91/127, 71.7%) (“I remember when I used to be addicted to Newport cigs now the smell of cigarette smoke gives me a headache #happyiquit #cigsarenasty”). However, almost 30% (36/127, 28.4%) of former smokers were positive or ambivalent about menthol (“even though i quit smoking im tempted to try some marlboro menthol lights, i heard they stink less and are less harmful”).

Figure 2. Positive and negative sentiment toward menthol cigarettes by domains and categories (n=6657). LLC: little cigars and cigarillos. *Pearson chi-square comparison between positive and negative sentiment, significant at *P*<.05.

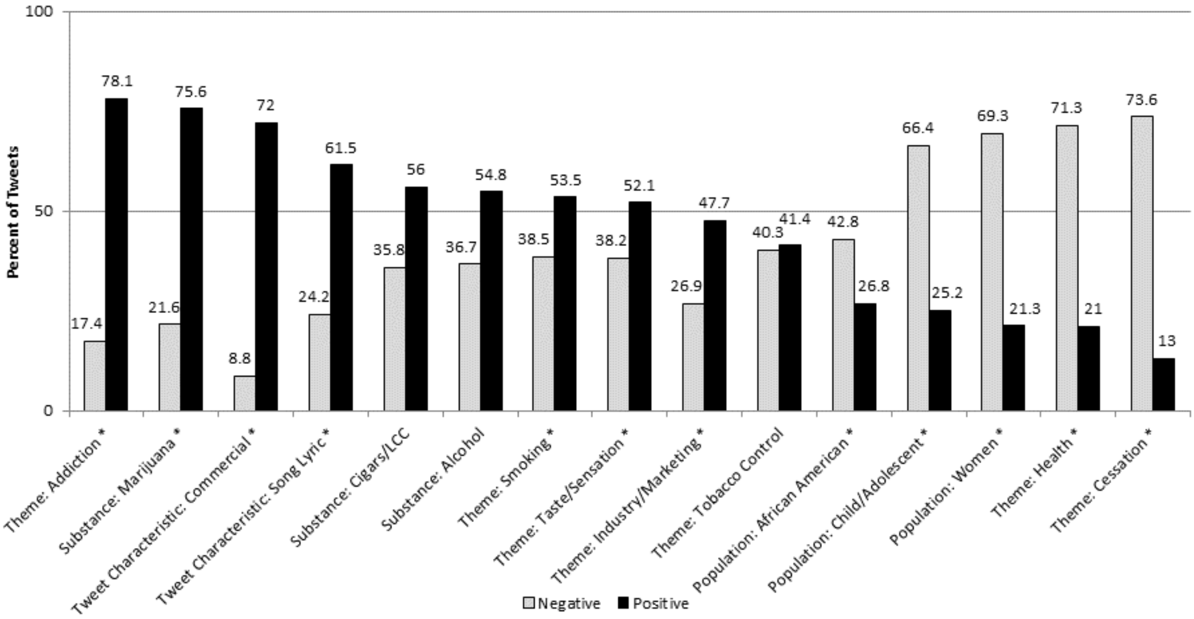
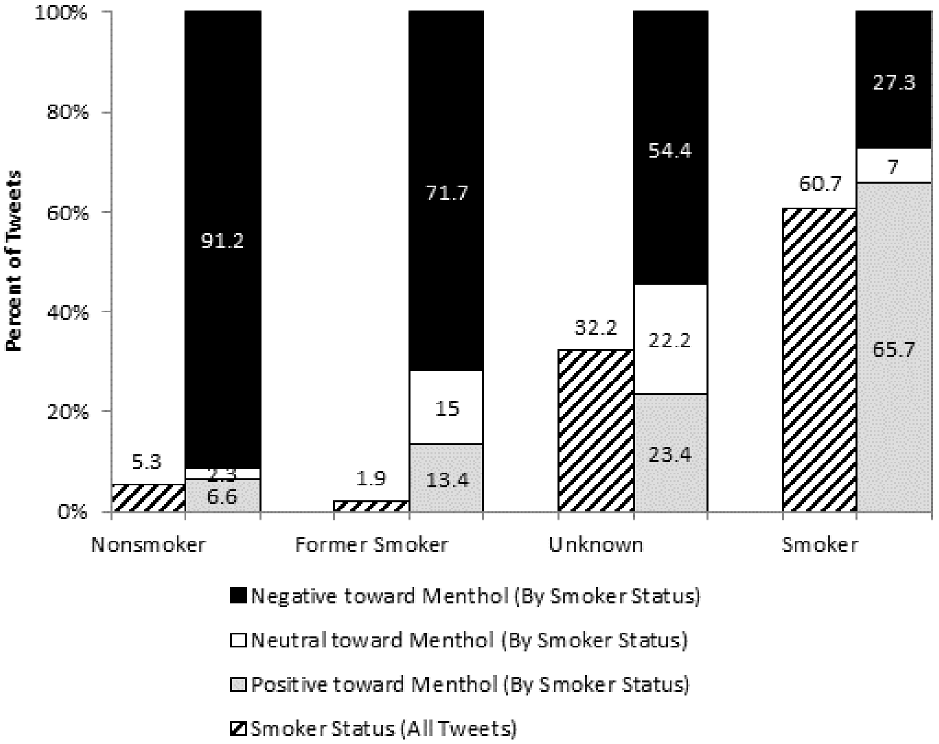


Figure 3. Tweets by smoker status and sentiment (n=6657).



Discussion

Principal Findings

Tweets about menthol cigarettes were most frequently about themes of smoking, taste or sensation, and addiction; African American populations; marijuana as an other substance; or song lyrics or pop culture references. Menthol content on Twitter was also driven by smokers who were more positive than others about menthol cigarettes. A prior study of tobacco-relevant content on Twitter found that most tweets were about personal experiences of use [34]. Liking or disliking the taste or sensation of menthol cigarettes found in a quarter of the tweets and in prior research [18,23] are understandable reasons for smoking menthol rather than nonmenthol cigarettes.

Additionally, Twitter users, particularly smokers, frequently mentioned marijuana and menthol cigarettes, and such tweets tended to be positive. This type of linkage has been found related to use of these products, with adolescent menthol smokers more likely to use marijuana in the past 30 days than nonmenthol smokers [49,50]. Similar to the effect of Twitter content on perceptions of the use of hookah [37] and marijuana [51], these positive views of menthol could normalize menthol cigarette use among the Twitter followers of these users.

Discussion over Twitter focused on smokers' concerns and tended to dwarf public health concerns, such as smoking cessation, health, and tobacco control. Some conversations about menthol cigarettes were cessation related but, as with other Twitter conversations about cessation, they lacked discussion of evidence-based strategies [52]. As in prior work, a small number of menthol smokers on Twitter may have perceived that switching to a nonmenthol cigarette is a good way to quit [26]. As in prior studies [18,53,54], we found that some smokers saw a medicinal effect of menthol cigarettes when they were sick. Studies have found that African American smokers may perceive health benefits of menthol or view menthol cigarettes as less harmful than nonmenthol cigarettes—beliefs that may interfere with quitting [14,18,20,23,24,26]. However, this view of menthol cigarettes as less harmful was not prevalent on Twitter and may reflect differences in broader population views of menthol. For instance, a national study of US smokers and nonsmokers found that 45.8% of adults believed that menthol and nonmenthol cigarettes were equally harmful; however, a sizable minority did not know whether menthol cigarettes were more or less harmful than nonmenthol cigarettes (40.8%) [21]. Few conversations on Twitter reflected concern for tobacco control policies. There was limited discussion about a menthol cigarette ban in the EU. After the TPSAC report in 2011 [12], but before an advanced notice of proposed rulemaking on menthol in cigarettes in 2013 [55], there was almost no Twitter discussion about a possible menthol cigarette ban in the United States or any other jurisdiction.

Finally, tweets reflected linkage of menthol cigarettes with African American populations historically targeted by menthol marketing and that have higher use prevalence [4,56]. Prior work has identified this "African Americanization" through targeted marketing and financial support for African American organizations [56]. African American menthol smokers have recognized that they have been targeted by menthol cigarette advertising [26]. This linkage between menthol cigarettes and African Americans is also seen on Twitter and is characterized by negative sentiment toward menthol cigarettes. Unfortunately, this negative sentiment seems to be driven by negative stereotypes about African American smokers rather than by rejection of the targeted marketing of menthol cigarettes to African Americans.

Strengths and Limitations

The strengths of this study are that it included a full year of global Twitter data from the Firehose, representing the entire corpus, rather than a sampling of Twitter content. Limitations are that, because we drew menthol content only from existing smoking and cigarette datasets, we may have missed menthol cigarette-relevant Twitter content without those terms. Additionally, due to this limitation, our research cannot be generalized to all tweets. We could not use Kool and Salem brand names as keywords because they were not precise enough to characterize menthol cigarette-relevant conversations. Future research in this area should consider using the menthol cigarette keywords beyond these existing archives and also examine potential substitution of menthol cigarettes with menthol flavors of other tobacco products such as e-cigarettes or cigars. The study lacked information about demographic or geographic information of Twitter users to understand differences in perceptions. However, themes found in prior research about menthol cigarettes [18-21,43] were also found in this study, suggesting these perceptions may be shared broadly. Future research can focus more in-depth on discussion of menthol cigarettes related to different populations. Another limitation was the cross-sectional nature of our analysis. Though the analysis was appropriate given our interest in the overall frequency of content about menthol cigarettes in this time period, a longitudinal analysis could provide information about changes in perceptions of menthol cigarette Twitter content.

Conclusion

Our findings present a relatively new way of assessing public opinions of menthol cigarettes through Twitter messages. Most messages are generated by smokers who have more positive sentiment toward these harmful combustible products. Misperceptions of menthol cigarettes having medicinal effects are prevalent, and positive linkage with marijuana is common. Addressing these common misconceptions and denormalizing menthol cigarette use, particularly for African American smokers, could ultimately save lives. In the United States, a menthol cigarette ban would be estimated to avert over 600,000 deaths by 2050, almost one-third among African Americans [57]. Framing public communication about menthol cigarettes to denormalize use in both online and offline forums is especially critical in light of potential regulation by the FDA, the EU, other jurisdictions, and localities. Future analyses could also use Twitter to examine reaction to regulations restricting menthol cigarettes among likely smokers and nonsmokers.

Acknowledgments

SWR conceived of research leading to this paper, conducted the analysis, and drafted the manuscript. SWR, CLJ, SB, and MB coded the tweets. SE leads the parent study and KMR provided oversight for the study. All authors made significant contributions to article conceptualization and editing and approved the final draft.

Funding for this study was provided by *Tobacco Control in a Rapidly Changing Media Environment* (U01CA154254) from the National Cancer Institute’s State and Community Tobacco Control Initiative. Additional support was provided by Truth Initiative for SWR’s time and the National Institute On Drug Abuse of the National Institutes of Health for CLJ’s effort (Award Number F31DA039609). 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

During the preparation of this manuscript, KMR was a member of the TPSAC for the FDA Center for Tobacco Products. The views expressed in this paper are his and not those of the FDA. All other authors have no competing interests.

Multimedia Appendix 1

Detailed study codebook.

[PDF File (Adobe PDF File), 501KB - jmir_v19i2e56_app1.pdf]

References

1. Carpenter CM, Wayne GF, Connolly GN. The role of sensory perception in the development and targeting of tobacco products. Addiction 2007 Jan;102(1):136-147. [doi: 10.1111/j.1360-0443.2006.01649.x] [Medline: 17207131]
2. Dessirier JM, O'Mahony M, Carstens E. Oral irritant properties of menthol: sensitizing and desensitizing effects of repeated application and cross-desensitization to nicotine. Physiol Behav 2001 May;73(1-2):25-36. [Medline: 11399291]
3. Federal Trade Commission. Federal Trade Commission Cigarette Report for 2012. Washington, DC: US Federal Trade Commission; 2015.
4. Caraballo RS, Asman K. Epidemiology of menthol cigarette use in the United States. Tob Induc Dis 2011;9 Suppl 1:S1 [FREE Full text] [doi: 10.1186/1617-9625-9-S1-S1] [Medline: 21624147]
5. Fallin A, Goodin AJ, King BA. Menthol cigarette smoking among lesbian, gay, bisexual, and transgender adults. Am J Prev Med 2015 Jan;48(1):93-97 [FREE Full text] [doi: 10.1016/j.amepre.2014.07.044] [Medline: 25245795]

6. The NSDUH Report: Use of Menthol Cigarettes. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies; Nov 19, 2009.
7. Giovino GA, Sidney S, Gfroerer JC, O'Malley PM, Allen JA, Richter PA, et al. Epidemiology of menthol cigarette use. Nicotine Tob Res 2004 Feb;6 Suppl 1:S67-S81. [doi: 10.1080/14622203710001649696] [Medline: 14982710]
8. Giovino GA, Villanti AC, Mowery PD, Sevilimedu V, Niaura RS, Vallone DM, et al. Differential trends in cigarette smoking in the USA: is menthol slowing progress? Tob Control 2015 Jan;24(1):28-37. [doi: 10.1136/tobaccocontrol-2013-051159] [Medline: 23997070]
9. Anderson SJ. Marketing of menthol cigarettes and consumer perceptions: a review of tobacco industry documents. Tob Control 2011 May;20 Suppl 2:ii20-ii28 [FREE Full text] [doi: 10.1136/tc.2010.041939] [Medline: 21504928]
10. Washington HA. Burning love: big tobacco takes aim at LGBT youths. Am J Public Health 2002 Jul;92(7):1086-1095. [Medline: 12084686]
11. Family Smoking Prevention and Tobacco Control Act of 2009, HR 1256, 111th Cong, 1st Sess. 2009. URL: https://www.gpo.gov/fdsys/pkg/PLAW-111publ31/pdf/PLAW-111publ31.pdf [accessed 2017-02-08] [WebCite Cache ID 6o8A5acIK]
12. Tobacco Products Scientific Advisory Committee. Menthol Cigarettes and Public Health: Review of the Scientific Evidence and Recommendations. Silver Spring, MD: U.S. Food and Drug Administration; 2011. URL: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/TobaccoProductsScientificAdvisoryCommittee/UCM269697.pdf [accessed 2017-02-08] [WebCite Cache ID 6o89Eo5qG]
13. Food and Drug Administration. Preliminary scientific evaluation of the possible public health effects of menthol versus nonmenthol cigarettes. Silver Spring, MD: FDA; 2013. URL: http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/PeerReviewofScientificInformationandAssessments/UCM361598.pdf [accessed 2017-02-08] [WebCite Cache ID 6o89yeyhd]
14. Benowitz NL, Samet JM. The threat of menthol cigarettes to U.S. public health. N Engl J Med 2011 Jun 9;364(23):2179-2181. [doi: 10.1056/NEJMp1103610] [Medline: 21542737]
15. Banks E, Joshy G, Weber MF, Liu B, Grenfell R, Egger S, et al. Tobacco smoking and all-cause mortality in a large Australian cohort study: findings from a mature epidemic with current low smoking prevalence. BMC Med 2015;13:38 [FREE Full text] [doi: 10.1186/s12916-015-0281-z] [Medline: 25857449]
16. Siegel M. A lost opportunity for public health--the FDA advisory committee report on menthol. N Engl J Med 2011 Jun 9;364(23):2177-2179. [doi: 10.1056/NEJMp1103403] [Medline: 21542736]
17. Framework Convention Alliance. Canadian provinces first in world to ban menthol flavouring. 2015 Nov 16. URL: http://www.fctc.org/fca-news/opinion-pieces/1328-canadian-provinces-first-to-ban-menthol-flavouring [accessed 2016-02-24] [WebCite Cache ID 6fXdhUP9B]
18. Allen B, Cruz TB, Leonard E, Unger JB. Development and validation of a scale to assess attitudes and beliefs about menthol cigarettes among african american smokers. Eval Health Prof 2010 Dec;33(4):414-436. [doi: 10.1177/0163278710383276] [Medline: 21149394]
19. Wackowski OA, Delnevo CD, Lewis MJ. Risk perceptions of menthol cigarettes compared with nonmenthol cigarettes among New Jersey adults. Nicotine Tob Res 2010 Jul;12(7):786-790. [doi: 10.1093/ntr/ntq085] [Medline: 20522521]
20. Castro FG. Physiological, psychological, social, and cultural influences on the use of menthol cigarettes among Blacks and Hispanics. Nicotine Tob Res 2004 Feb;6 Suppl 1:S29-S41. [doi: 10.1080/14622200310001649487] [Medline: 14982707]
21. Davis SP, McClave-Regan AK, Rock VJ, Kruger J, Garrett BE. Perceptions of menthol cigarette use among U.S. adults and adult smokers: findings from the 2009 HealthStyles survey. Nicotine Tob Res 2010 Dec;12 Suppl 2:S125-S135. [doi: 10.1093/ntr/ntq205] [Medline: 21177369]
22. Brennan E, Gibson L, Momjian A, Hornik RC. Are young people's beliefs about menthol cigarettes associated with smoking-related intentions and behaviors? Nicotine Tob Res 2015 Jan;17(1):81-90 [FREE Full text] [doi: 10.1093/ntr/ntu134] [Medline: 25151661]
23. Unger JB, Allen B, Leonard E, Wenten M, Cruz TB. Menthol and non-menthol cigarette use among Black smokers in Southern California. Nicotine Tob Res 2010 Apr;12(4):398-407. [doi: 10.1093/ntr/ntq016] [Medline: 20167636]
24. Allen B, Unger JB. Sociocultural correlates of menthol cigarette smoking among adult African Americans in Los Angeles. Nicotine Tob Res 2007 Apr;9(4):447-451. [doi: 10.1080/14622200701239647] [Medline: 17454699]
25. Kreslake JM, Wayne GF, Connolly GN. The menthol smoker: tobacco industry research on consumer sensory perception of menthol cigarettes and its role in smoking behavior. Nicotine Tob Res 2008 Apr;10(4):705-715. [doi: 10.1080/14622200801979134] [Medline: 18418792]
26. Richter P, Beistle D, Pederson L, O'Hegarty M. Small-group discussions on menthol cigarettes: listening to adult African American smokers in Atlanta, Georgia. Ethn Health 2008 Apr;13(2):171-182. [doi: 10.1080/13557850701784694] [Medline: 18425713]
27. Donelle L, Booth RG. Health tweets: an exploration of health promotion on twitter. Online J Issues Nurs 2012 Sep;17(3):4 [FREE Full text] [Medline: 23036060]
28. Lyles CR, López A, Pasick R, Sarkar U. “5 mins of uncomfyness is better than dealing with cancer 4 a lifetime”: an exploratory qualitative analysis of cervical and breast cancer screening dialogue on Twitter. J Cancer Educ 2013 Mar;28(1):127-133. [doi: 10.1007/s13187-012-0432-2] [Medline: 23132231]

29. Terdiman D. CNET news. 2012 Oct 26. Report: Twitter hits half a billion tweets a day URL: http://www.cnet.com/news/report-twitter-hits-half-a-billion-tweets-a-day/ [accessed 2016-02-24] [WebCite Cache ID 6fXdue7b7]
30. Duggan M. Mobile messaging and social media. Washington, DC: Pew Research Center; 2015. URL: http://www.pewinternet.org/files/2015/08/Social-Media-Update-2015-FINAL2.pdf [accessed 2017-02-09] [WebCite Cache ID 6o9KMjyaC]
31. Statista. Leading social networks worldwide as of August 2015, ranked by number of active users (in millions). 2015 Nov 16. URL: http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [accessed 2016-02-24] [WebCite Cache ID 6fXe06C1d]
32. Huang J, Kornfield R, Szczypka G, Emery SL. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control 2014 Jul;23 Suppl 3:iii26-iii30 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051551] [Medline: 24935894]
33. Kim AE, Hopper T, Simpson S, Nonnemaker J, Lieberman AJ, Hansen H, et al. Using Twitter data to gain insights into e-cigarette marketing and locations of use: an infoveillance study. J Med Internet Res 2015;17(10):e251 [FREE Full text] [doi: 10.2196/jmir.4466] [Medline: 26545927]
34. Myslín M, Zhu S, Chapman W, Conway M. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
35. Cole-Lewis H, Pugatch J, Sanders A, Varghese A, Posada S, Yun C, et al. Social listening: a content analysis of e-cigarette discussions on Twitter. J Med Internet Res 2015;17(10):e243 [FREE Full text] [doi: 10.2196/jmir.4969] [Medline: 26508089]
36. Jo CL, Kornfield R, Kim Y, Emery S, Ribisl KM. Price-related promotions for tobacco products on Twitter. Tob Control 2016 Jul;25(4):476-479. [doi: 10.1136/tobaccocontrol-2015-052260] [Medline: 26130449]
37. Krauss MJ, Sowles SJ, Moreno M, Zewdie K, Grucza RA, Bierut LJ, et al. Hookah-related Twitter chatter: a content analysis. Prev Chronic Dis 2015;12:E121 [FREE Full text] [doi: 10.5888/pcd12.150140] [Medline: 26226068]
38. McDaniel PA, Patzke H, Malone RE. Twitter users' reaction to a chain pharmacy's decision to end tobacco sales. Tob Induc Dis 2015;13:36 [FREE Full text] [doi: 10.1186/s12971-015-0060-9] [Medline: 26539069]
39. Internet Live Stats. Twitter usage statistics.: InternetLiveStats.com URL: http://www.internetlivestats.com/twitter-statistics/#trend [accessed 2017-02-09] [WebCite Cache ID 6o9LT01k6]
40. Health Media Collaboratory. URL: http://www.healthmediacollaboratory.org [accessed 2016-02-24] [WebCite Cache ID 6fXYgaG3T]
41. Gnip. Boulder, CO: Gnip, Inc; 2016. URL: https://gnip.com/ [accessed 2016-02-24] [WebCite Cache ID 6fXdIvOvw]
42. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
43. Winickoff JP, McMillen RC, Vallone DM, Pearson JL, Tanski SE, Dempsey JH, et al. US attitudes about banning menthol in cigarettes: results from a nationally representative survey. Am J Public Health 2011 Jul;101(7):1234-1236. [doi: 10.2105/AJPH.2011.300146] [Medline: 21566038]
44. Amonini C, Donovan RJ. The relationship between youth's moral and legal perceptions of alcohol, tobacco and marijuana and use of these substances. Health Educ Res 2006 Apr;21(2):276-286 [FREE Full text] [doi: 10.1093/her/cyh064] [Medline: 16234282]
45. Richter PA, Pederson LL, O'Hegarty MM. Young adult smoker risk perceptions of traditional cigarettes and nontraditional tobacco products. Am J Health Behav 2006;30(3):302-312. [doi: 10.5555/ajhb.2006.30.3.302] [Medline: 16712444]
46. Kim Y, Huang J, Emery S. Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. J Med Internet Res 2016;18(2):e41 [FREE Full text] [doi: 10.2196/jmir.4738] [Medline: 26920122]
47. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. Am J Prev Med 2008 Aug;35(2):158-176 [FREE Full text] [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
48. Emery SL, Szczypka G, Abril EP, Kim Y, Vera L. Are you scared yet? Evaluating fear appeal messages in tweets about the tips campaign. J Commun 2014 Apr;64:278-295 [FREE Full text] [doi: 10.1111/jcom.12083] [Medline: 25429162]
49. Kong G, Singh N, Camenga D, Cavallo D, Krishnan-Sarin S. Menthol cigarette and marijuana use among adolescents. Nicotine Tob Res 2013 Dec;15(12):2094-2099 [FREE Full text] [doi: 10.1093/ntr/ntt102] [Medline: 23884319]
50. Azagba S, Sharaf MF. Binge drinking and marijuana use among menthol and non-menthol adolescent smokers: findings from the youth smoking survey. Addict Behav 2014 Mar;39(3):740-743. [doi: 10.1016/j.addbeh.2013.12.005] [Medline: 24369112]
51. 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. [doi: 10.1016/j.jadohealth.2014.10.270] [Medline: 25620299]
52. Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. Tob Control 2012 Jul;21(4):447-449 [FREE Full text] [doi: 10.1136/tc.2010.042507] [Medline: 21730101]
53. Klausner K. Menthol cigarettes and smoking initiation: a tobacco industry perspective. Tob Control 2011 May;20 Suppl 2:ii12-ii19 [FREE Full text] [doi: 10.1136/tc.2010.041954] [Medline: 21504927]

54. Truth Tobacco Documents Library. A study of smokers' habits and attitudes with special emphasis on low tar and menthol cigarettes, volume I. Bates No:2040500265-2040500399. 1979. URL: https://industrydocuments.library.ucsf.edu/tobacco/docs/ [accessed 2016-02-24] [WebCite Cache ID 6fXe95h1b]
55. Food and Drug Administration, Department of Health and Human Services. Menthol in cigarettes, tobacco products; request for comments. 21 CFR Part 1140 [Docket No. FDA-2013-N-0521]. Federal Register 2013 Jul 24;78(142):44484-44485 [FREE Full text]
56. Gardiner PS. The African Americanization of menthol cigarette use in the United States. Nicotine Tob Res 2004 Feb;6 Suppl 1:S55-S65. [doi: 10.1080/14622200310001649478] [Medline: 14982709]
57. Levy DT, Pearson JL, Villanti AC, Blackman K, Vallone DM, Niaura RS, et al. Modeling the future effects of a menthol ban on smoking prevalence and smoking-attributable deaths in the United States. Am J Public Health 2011 Jul;101(7):1236-1240. [doi: 10.2105/AJPH.2011.300179] [Medline: 21566034]

Abbreviations

EU: European Union
FDA: Food and Drug Administration
HMC: Health Media Collaboratory
LGBT: lesbian, gay, bisexual, or transgender
TPSAC: Tobacco Products Science Advisory Committee

Edited by G Eysenbach; submitted 24.02.16; peer-reviewed by S Dunlop, M Gonzalez, P Shafer; comments to author 27.06.16; revised version received 21.08.16; accepted 30.01.17; published 27.02.17.

Please cite as:
Rose SW, Jo CL, Binns S, Buenger M, Emery S, Ribisl KM
Perceptions of Menthol Cigarettes Among Twitter Users: Content and Sentiment Analysis
J Med Internet Res 2017;19(2):e56
URL: http://www.jmir.org/2017/2/e56/
doi:10.2196/jmir.5694
PMID:28242592

Original Paper

Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study

Danielle Mowery[1], MS, PhD; Hilary Smith[2], BS; Tyler Cheney[2], BS; Greg Stoddard[3], MS; Glen Coppersmith[4,5], MA, PhD; Craig Bryan[2], PsyD, ABPP; Mike Conway[1], PhD

[1]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States
[2]Department of Psychology, University of Utah, Salt Lake City, UT, United States
[3]Department of Family And Preventive Medicine, University of Utah, Salt Lake City, UT, United States
[4]Qntfy, Crownsville, MD, United States
[5]Human Language Technology Center of Excellence, John Hopkins University, Baltimore, MD, United States

Corresponding Author:
Danielle Mowery, MS, PhD
Department of Biomedical Informatics
University of Utah
421 Wakara Way, Ste 140
Salt Lake City, UT, 84108
United States
Phone: 1 8015856739
Fax: 1 (801) 581 4297
Email: danielle.mowery@utah.edu

Abstract

Background: With a lifetime prevalence of 16.2%, major depressive disorder is the fifth biggest contributor to the disease burden in the United States.

Objective: The aim of this study, building on previous work qualitatively analyzing depression-related Twitter data, was to describe the development of a comprehensive annotation scheme (ie, coding scheme) for manually annotating Twitter data with Diagnostic and Statistical Manual of Mental Disorders, Edition 5 (DSM 5) major depressive symptoms (eg, depressed mood, weight change, psychomotor agitation, or retardation) and Diagnostic and Statistical Manual of Mental Disorders, Edition IV (DSM-IV) psychosocial stressors (eg, educational problems, problems with primary support group, housing problems).

Methods: Using this annotation scheme, we developed an annotated corpus, Depressive Symptom and Psychosocial Stressors Acquired Depression, the SAD corpus, consisting of 9300 tweets randomly sampled from the Twitter application programming interface (API) using depression-related keywords (eg, depressed, gloomy, grief). An analysis of our annotated corpus yielded several key results.

Results: First, 72.09% (6829/9473) of tweets containing relevant keywords were nonindicative of depressive symptoms (eg, “we’re in for a new economic depression”). Second, the most prevalent symptoms in our dataset were depressed mood and fatigue or loss of energy. Third, less than 2% of tweets contained more than one depression related category (eg, diminished ability to think or concentrate, depressed mood). Finally, we found very high positive correlations between some depression-related symptoms in our annotated dataset (eg, fatigue or loss of energy and educational problems; educational problems and diminished ability to think).

Conclusions: We successfully developed an annotation scheme and an annotated corpus, the SAD corpus, consisting of 9300 tweets randomly-selected from the Twitter application programming interface using depression-related keywords. Our analyses suggest that keyword queries alone might not be suitable for public health monitoring because context can change the meaning of keyword in a statement. However, postprocessing approaches could be useful for reducing the noise and improving the signal needed to detect depression symptoms using social media.

(J Med Internet Res 2017;19(2):e48) doi:10.2196/jmir.6895

KEYWORDS

social media; Twitter messaging; natural language processing; major depressive disorder; data annotation; machine learning

Introduction

Background

With a lifetime prevalence of 16.2% in the United States [1], major depressive disorder is the fifth biggest contributor to the disease burden in the United States [2]. Several national face-to-face and telephonic interview-based surveys in the United States aim to better understand the prevalence of depressive symptoms in the community. However, these surveys are both episodic and expensive to conduct. Social media platforms like Twitter, in conjunction with "big data" technologies like natural language processing and machine learning, support processing very large datasets and may provide a scalable means of both monitoring depressive disorder over time and providing new insights to better our understanding of depression (and mental illness more generally). As part of our goal of developing language technologies capable of accurately identifying depressive symptoms, we have developed a large manually annotated (coded) corpus or collection of Twitter posts (tweets) coded according to depressive symptoms and psychosocial stressors derived primarily from Diagnostic and Statistical Manual of Mental Disorders, Edition 5 (DSM 5; depressive symptoms) [3] and DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Edition IV (DSM-IV Axis IV; psychosocial stressors) [4]. This annotated corpus allows us to better understand the language used to express depressive symptoms and psychosocial stressors associated with depression, to identify relationships between depressive symptoms and psychosocial stressors expressed in tweets, and ultimately, to facilitate the development of a natural language processing system capable of automatically identifying depressive symptoms and psychosocial stressors from Twitter data.

Social Media

The use of social media for health applications, particularly in the public health domain, is a rapidly growing area of research [5,6]. For example, social media has been leveraged to monitor infectious disease outbreaks [7,8] and understand prescription drug and smoking behaviors [9-11]. The value of social media for understanding mental health is particularly marked, given that it provides—in the case of Twitter—access to public, first person accounts of user behaviors, activities, thoughts, and feelings that may be indicative of emotional well-being [12]. Twitter in particular has several advantages as a resource for data. First, as of August 2015, Twitter has been used by 23% of adults in the United States, with slightly more men (25%) than women (21%) using the service [13]. Second, Twitter is a "broadcast" social network, with all the data public by default. Third, acquiring Twitter data via the free public application programming interface (API) or commercial data resellers (eg, *gnip* [14]) is a relatively straightforward process. However, the use of Twitter data does present a number of challenges. First, the brevity of Twitter posts (≤140 characters) frequently provides insufficient context to confidently interpret a post. Second, the informal nature of the language used in social media posts (eg, "tiredddd") means that natural language processing techniques and tools developed for more formal texts are likely to perform less well on Twitter data [15]. Third, Twitter posts often exhibit creative spellings and missing spaces (eg, "sodepressed"), presenting challenges for automatic processing. Finally, Twitter users may selectively discuss topics of interest with researchers; for example, some individuals may not feel comfortable discussing disease-related symptoms on social media due to concerns of privacy and stigmatization [16].

Major Depressive Disorder

The American Psychiatric Association defines major depressive disorder as continuously experiencing *depressed mood* and *anhedonia* for 2 weeks or more as well as one or more of the following symptoms: *fatigue*, *inappropriate guilt*, *difficulty concentrating*, *psychomotor agitation or retardation*, or *weight loss or gain* [3,4]. These symptoms make major depressive disorder one of the most debilitating and burdensome global diseases [17,18], with an economic impact estimated to be US $2.5 trillion in 2010 [19]. For individuals living with depression, the disorder can substantially reduce quality of life in several areas, including interactions with others, productivity at work, and quality of sleep and nutrition [20]. Depression has also been correlated with other high-risk behaviors and chronic diseases, including smoking [21], alcohol consumption [22], physical inactivity [23], and sleep disturbance [20,24].

Population-Level Depression Surveys

Given the range and extent that depression affects a given population, several surveys, programs, and diagnostic tools have been developed to better understand or diagnose depressive disorder. For example, in the United States, the *National Survey on Drug Use and Health (NSDUH)* provides national, state, and local data related to alcohol, tobacco, illegal drug use and abuse, and mental disorders, including nonincarcerated citizens of age 12 and older [25]. The *Youth Risk Behavior Surveillance System (YRBSS)* monitors behaviors such as alcohol and other drug use, tobacco use, and unhealthy dietary behaviors, and so on, and their correspondence with death and disability among youth and adults [26]. *The Behavioral Risk Factors Surveillance System (BRFSS)* is a telephone survey that collects data from across the United States, including health-related risk behaviors, chronic health conditions, and use of preventive services [27]. The *BRFSS - Anxiety and Depression Optional Module* specifically collects information at the state level to assess the prevalence of anxiety and depressive disorders with questions that closely mirror the DSM 5 major depression criteria.

Related Works

Major Depressive Disorder and Social Media

Recent work at the intersection of computer science, public health, and psychology suggests that social media can be leveraged to better understand, identify, and characterize depression [12]. For example, De Choudhury et al used a crowdsourcing data generation method in conjunction with machine learning to identify depression-indicative tweets at scale [28], whilst a follow-up study investigated the characteristics of Twitter users prior to the onset of depression, discovering that *decrease in social activity*, *raised negative affect*, *highly clustered ego networks*, *heightened relational and medical concerns*, and *greater expression of religious* *involvement* were all characteristic of the onset of depression [29].

In a study using Facebook, Schwartz et al used status updates and personality survey results as features in a regression model to classify the degree of depression of 28,749 Facebook users [30]. A temporal analysis of these posts demonstrated that mood worsens in the transition from summer to winter for users. Coppersmith et al further characterized the language of mental illnesses (eg, depression) by identifying tweets containing self-reported diagnosis ("I was diagnosed with depression today"), then analyzing the linguistic characteristics of tweets from both a self-reported depression and a control group, observing that the usage of words from the Linguistic Inquiry and Word Count (LIWC) lexicon [31] associated with negative emotions including *anxiety* and *anger*, biological states such as *health* and *death*, cognitive mechanisms including *cause* and *tentativeness*, and syntactic usage of *first person pronoun* (eg, "I") may distinguish a depressed from a nondepressed individual [32,33]. Preotuic-Pietro et al observed many features that distinguish depressed Twitter users from controls [34], for example, terms associated with *illness management* (eg, "meds," "pills," and "therapy") and *increased focus on the self* (eg, "I," "I am," "I have," "I was," and "myself").

In this study, we build on these existing efforts by developing an annotation scheme for encoding depressive symptoms and psychosocial stressors associated with major depressive disorder in Twitter tweets and conducting analyses to provide insights into how users express these symptoms on Twitter. From these analyses, specifically, we aim to (1) validate the annotation scheme, (2) learn the predictive value of depression-related keywords with respect to identifying depressive symptoms and psychosocial stressors, (3) determine the frequency of depressive symptoms and psychosocial stressors expressed, (4) learn new predictive words for each depressive symptom and psychosocial stressor, and (5) assess whether particular depressive symptoms and psychosocial stressors are correlated with one another.

Methods

Developing a Depression Annotation Scheme and Corpus for Twitter

All the data were collected from the Twitter API complying with Twitter's terms of service.

Developing an Annotation Scheme

In order to understand the various ways indicators of major depression disorder could be expressed in tweets and address our goal of building a dataset that can be used to train and test machine learning algorithms, we developed an annotation scheme (coding scheme) based on 6 resources:

Depression symptoms as described in the *Diagnostic and Statistical Manual of Mental Disorders, Edition 5 (DSM-V)* [3];

Psychosocial stressors described in *Axis IV of the Diagnostic and Statistical Manual of Mental Disorders, Edition IV (DSM-IV)* [4];

Depression symptoms described in the *Behavioral Risk Factors Surveillance System—Depression Module* [27];

Depression symptoms described in the *Harvard Department of Psychiatry National Depression Screening Day Scale (HANDS)* [35];

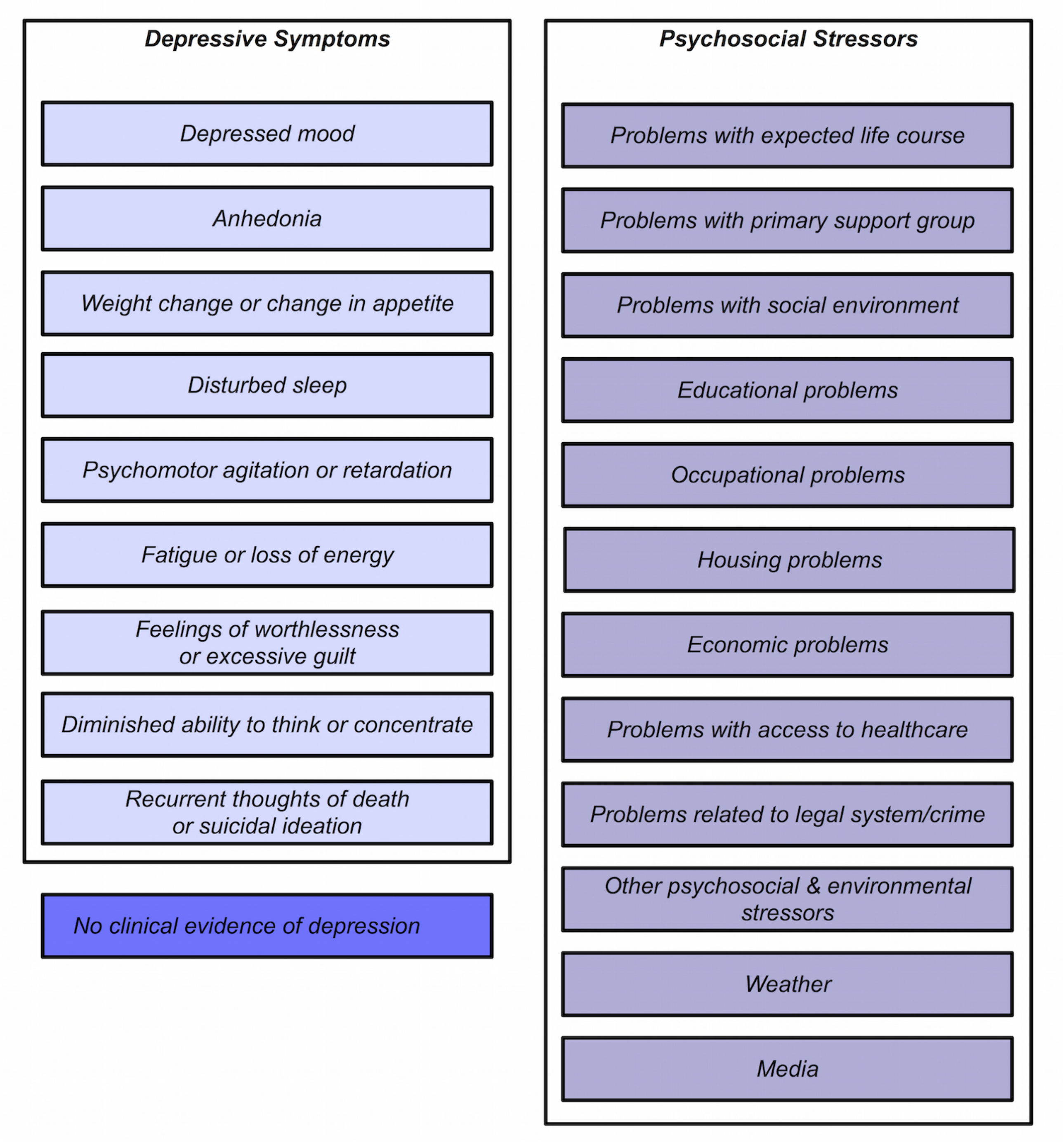
Depression symptoms described in the *Patient Health Questionnaire (PHQ-9)* [36];

Depression symptoms described in the *Quick Inventory of Depressive Symptomatology (QIDS-SR)* [37]; and

Suicide risk factors derived from the *Columbia Suicide Severity Scale* [38].

Finally, we enriched the annotation scheme with additional depression-related categories observed frequently in the data (*weather* and *media*). The resulting scheme contains depression symptom categories (9 parent categories) and psychosocial stressor categories (12 parent categories; Figure 1) [39]. Before finalizing the annotation scheme, both a psychiatrist and a counseling psychologist provided feedback on the annotation categories chosen and annotation instructions.

Figure 1. Major depressive disorder scheme (parent categories).



Building a Depression-Related Twitter Corpus

Data for our depression-related Twitter corpus were collected in two distinct ways. First, for our primary corpus construction effort, we searched the Twitter API using depression-related terms (Depressive Symptom and Psychosocial Stressors Acquired Depression, *SAD*, corpus). Second, we sampled the data collected as part of the *2015 Computational Linguistics and Clinical Psychology* (CLPsych) Shared Task [40]. Both corpora are described in detail below.

Depressive Symptoms and Psychosocial Stressor Acquired Depression (SAD) Corpus

We randomly selected Twitter user tweets from March 1 to March 31, 2013, using the Twitter API. For each day in March 2013, we randomly selected 300 tweets containing one or more keywords from the LIWC lexicon (eg, “die,” “pain,” and “tired”). We used the LIWC “sad” category keyword list and augmented this list with several keywords selected by a board-certified clinical psychologist (author CB). For example, the presence of the keyword “insomnia” might be suggestive of the depression symptom *disturbed sleep*. A complete list of keywords and associated depression stressors and symptoms can be found in Table 1 (n=110 total keywords).

Table 1. Linguistic Inquiry and Word Count (LIWC) concepts and associated keywords[a].

Depression categories		Linguistic Inquiry and Word Count
Depressive symptoms		
	Depressed mood	pain
	Weight change or change in appetite	appetite
	Disturbed sleep	insomnia
	Psychomotor agitation or retardation	restless, jitter*, groggy, dazed
	Fatigue or loss of energy	tired
	Feelings of worthlessness or excessive inappropriate guilt	guilt*, burden
	Diminished ability to think or concentrate, indecisiveness	concentrat*, focus*, indeci*
	Recurrent thoughts of death, suicidal ideation	suicid*, kill
Psychosocial stressors		
	Problems with primary support group	death, die*, funeral, cremat*, divorc, abus*, neglect*
	Occupational problems	fired, unemploy*
	Housing problems	homeless*
	LIWC “sad” keyword	abandon*, ache*, aching, agoni*, alone, broke*, cried, cries, crushed, cry, damag*, defeat*, depress*, depriv*, despair*, devastat*, disadvantage*, disappoint*, discourag*, dishearten*, disillusion*, dissatisf*, doom*, dull*, empt*, fail*, fatigu*, flunk*, gloom*, grave*, grief, griev*, grim*, heartbr*, helpless*, homesick*, hopeless*, hurt*, inadequa*, inferior*, isolat*, lame*, lone*, longing*, lose, loser*, loses, losing, loss*, lost, melanchol*, miser*, mourn*, neglect*, overwhelm* pathetic*, pessimis*, piti*, pity* , regret*, reject*, remorse*, resign*, ruin*, sad, sobbed, sobbing, sobs, solemn*, sorrow*, suffer*, tears*, traged*, tragic* , unhapp*, unimportant, unsuccessful*, useless*, weep*, wept, whine*, whining, woe*, worthless*, yearn*

[a] Depressive symptom *anhedonia* and psychosocial stressors such as *problems with expected life course with respect to self, problems related to the social environment, educational problems, economic problems, problems with access to health care, problems related to the legal system and crime, other psychosocial and environmental problems, weather,* and *media* do not have associated keywords.

CLPsych Corpus

In addition to the SAD corpus, we sampled tweets from a large corpus of Twitter data developed for the *2015* CLPsych shared task [40]. In order to build this corpus, CLPsych researchers queried Twitter (via the public Twitter API) for users with a self-disclosed, publicly stated psychiatric diagnosis (eg, “I was diagnosed with having depression”), then collected all available tweets from that user. The corpus consisted of up to 3000 tweets from each of the 477 users, from which we randomly sampled 100 users with self-disclosed depression diagnosis from the CLPsych dataset, located the “self-diagnosis” tweet, then annotated the subsequent 10 tweets from that user using our annotation scheme.

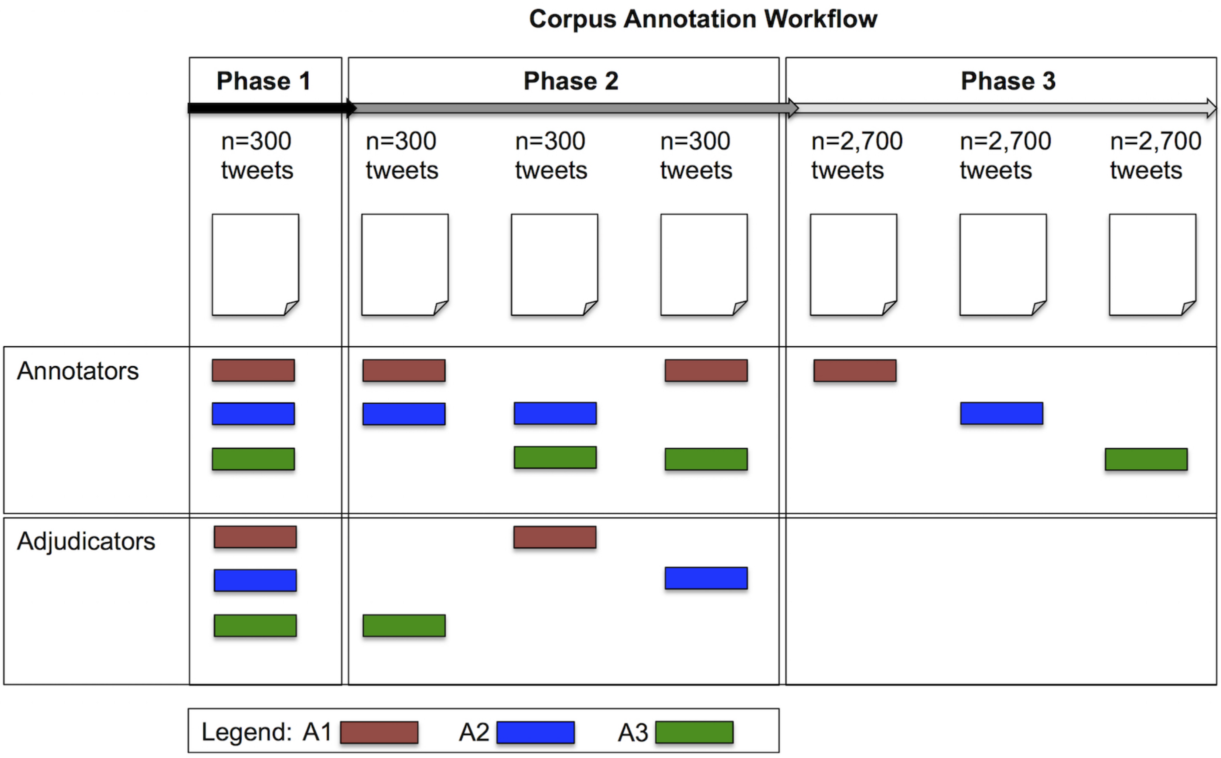
Validating the Annotation Scheme

In order to validate our annotation scheme, 3 annotators—2 psychology graduate researchers and a postdoctoral biomedical informatics researcher—annotated 1200 tweets from the SAD corpus in 3 phases. In phase 1, all 3 annotators annotated 300 tweets and reached agreement with consensus review. In phase 2, for the remaining 900 tweets and for all annotator pair combinations, 2 annotators independently annotated 300 tweets, and the remaining third annotator adjudicated any disagreements. For example, if annotators A1 and A2 annotated 300 tweets, annotator A3 would adjudicate those tweets where A1 and A2 disagreed regarding the appropriate label. We compared the annotations between each pair of annotators to determine the asserted categorical matches and mismatches. For example, a match occurs when both annotators (eg, A1 and A2) annotated the same category for the same tweet. There are 2 types of mismatches: type 1 mismatch occurs if A1 annotated a category for a tweet not annotated by A2; and a type 2 mismatch if A2 annotated a category for a tweet not annotated by A1. We report both overall and granular inter-annotator agreement between annotator pairs by comparing one annotator’s annotations to the other’s annotations (rather than assuming a ground truth) to compute *F* score [41]. *F* score is computed from the matches and mismatches and given as a percentage from the following equation:

$$F \text{ score}=(2\times\text{matches})/([2\times\text{matches}]+\text{mismatches}) \times 100\%$$

In phase 3, each annotator independently annotated 2700 tweets (8100 tweets total from 3 annotators) and to further ensure reliability, 1200 tweets were annotated by all 3 annotators. The resulting SAD corpus consists of 9300 tweets. A summary of this annotation workflow can be found in Figure 2. The CLPsych corpus was annotated by 1 annotator resulting in 1019 tweets (which are not included in the 9300 SAD tweets).

Figure 2. SAD corpus annotation in phases. A#=Annotator eg, A1=Annotator 1. SAD: Depressive Symptom and Psychosocial Stressors Acquired Depression.



Learning the Predictive Value of Depression-Related Keywords

For both the SAD and CLPsych corpora, in order to assess how accurately these depression-related keywords could identify depression-related tweets, we computed the precision of each depression-related keyword, defined as the count of tweets identified by the depression-related keyword and associated with a depression-related category divided by the total count of tweets identified by the depression-related keyword (tweet hits). For example, if 4 tweets were identified by the keyword “sobbing,” but only 1 of the 4 total tweets was encoded as a depressive symptom or psychosocial stressor, then the precision of the depression-related keyword is 25%. We classified the resulting precision using 5 equally sized categories:

1. zero to poor precision: 0-19%,
2. poor to low precision: 20-39%,
3. low to moderate precision: 40-59%,
4. moderate to high precision: 60-79%, and
5. high to excellent precision: 80-100%.

For each corpus and each precision category, we report the count of tweets identified by the count of depression-related keywords (tweet hits). Specifically, one or more keywords can match a single tweet, for example, the keywords “depressed” and “fired” in “I’m so depressed because I got fired today”; therefore, our denominator is the number of times a keyword was matched in a tweet.

Exploring the Frequency of Symptoms and Psychosocial Stressors

In order to estimate the proportion of said depressive symptoms and psychosocial stressors in our corpus, we characterized our total corpus of tweets by the proportion of tweets representing *no evidence of clinical depression* and *evidence of clinical depression*. Of the tweets representing evidence of clinical depression, we report the proportion of tweets representing depressive symptoms and psychosocial stressors. Finally, we provide example subtypes of depressive symptoms and psychosocial stressors. We compared the distributions of annotation categories between the SAD and CLPsych corpora in order to identify salient characteristics of Twitter users with a publicly stated diagnosis of depression.

Determining Predictive Word Features for Depressive Symptoms and Psychosocial Stressors

For both the SAD and CLPsych corpora, in order to identify words and phrases most characteristic of each category of depressive symptoms and psychosocial stressors (eg, the words most characteristic of, say, *occupational problem*), we used a technique referred to as feature selection [42] (keyword extraction in the corpus linguistics literature [43]). More specifically, we used the information gain metric [44] to compare the relative frequency of words associated with each depression category (eg, the word “fired” may appear more frequently in the *occupational problem* category than the *educational problem* category). The 10 most characteristic words—identified by information gain—are reported for each category. Specifically, we used Weka version 3.16.13 to learn words that occurred with the highest average rank for 5 independent subsets of the dataset [42].

Assessing Correlations Between Depressive Symptoms and Psychosocial Stressors

For the 9300 tweet SAD corpus only, in order to determine whether a correlation exists between 2 specific depressive symptoms and psychosocial stressors, we computed Pearson correlation coefficients for each pairwise combination of the 21 parent categories of depressive symptoms and psychosocial stressors from the annotation scheme. Given that each symptom or stressor category has only 2 states (annotated or not annotated), this correlation coefficient is sometimes called the phi coefficient, although the phi and Pearson correlation coefficients are algebraically identical. A higher correlation coefficient indicates that when the psychosocial stressor is annotated, the depressive symptom is more likely to also be annotated. We used the r value to interpret magnitude because P values are affected by sample size, whereas r values are not. We classified the correlation magnitude using Cohen effect size criteria into 4 categories [45]: less than small effect: <0.09; small to medium effect: 0.1-0.29; medium to large effect: 0.3-0.49; and greater than large effect: >0.50.

Results

Characterizing the Corpus

Our depression disorder scheme is comprised of 9 depressive symptoms and 12 psychosocial stressor categories that were applied to the SAD and CLPsych Twitter corpora. We observed an average number of 14-15 words with a standard deviation between 7 and 8 words (Table 2).

Table 2. Comparison of characteristics by corpus.

Characteristic	SAD	CLPsych
Query-level	tweet-level	user-level
Number of unique tweets	9300	1019
Number of unique words	19,822	3258
Average number of words per tweet (SD)	14.56 (7.40)	15.44 (8.07)

Validating the Annotation Scheme

We observed high overall interannotator agreement (F scores) between annotator pairs: ranging from 76% to 81% (Table 3). Overall F scores dropped slightly when comparing matches for all 3 annotators. Across pairs, we observed F scores ranging from 86% to 89% for *no evidence of clinical depression*. F scores varied widely across all annotated categories. High F scores were observed across annotator pairs for the depression symptom *fatigue or loss of energy* and psychosocial stressors *recurrent thoughts of death and suicidal ideation*.

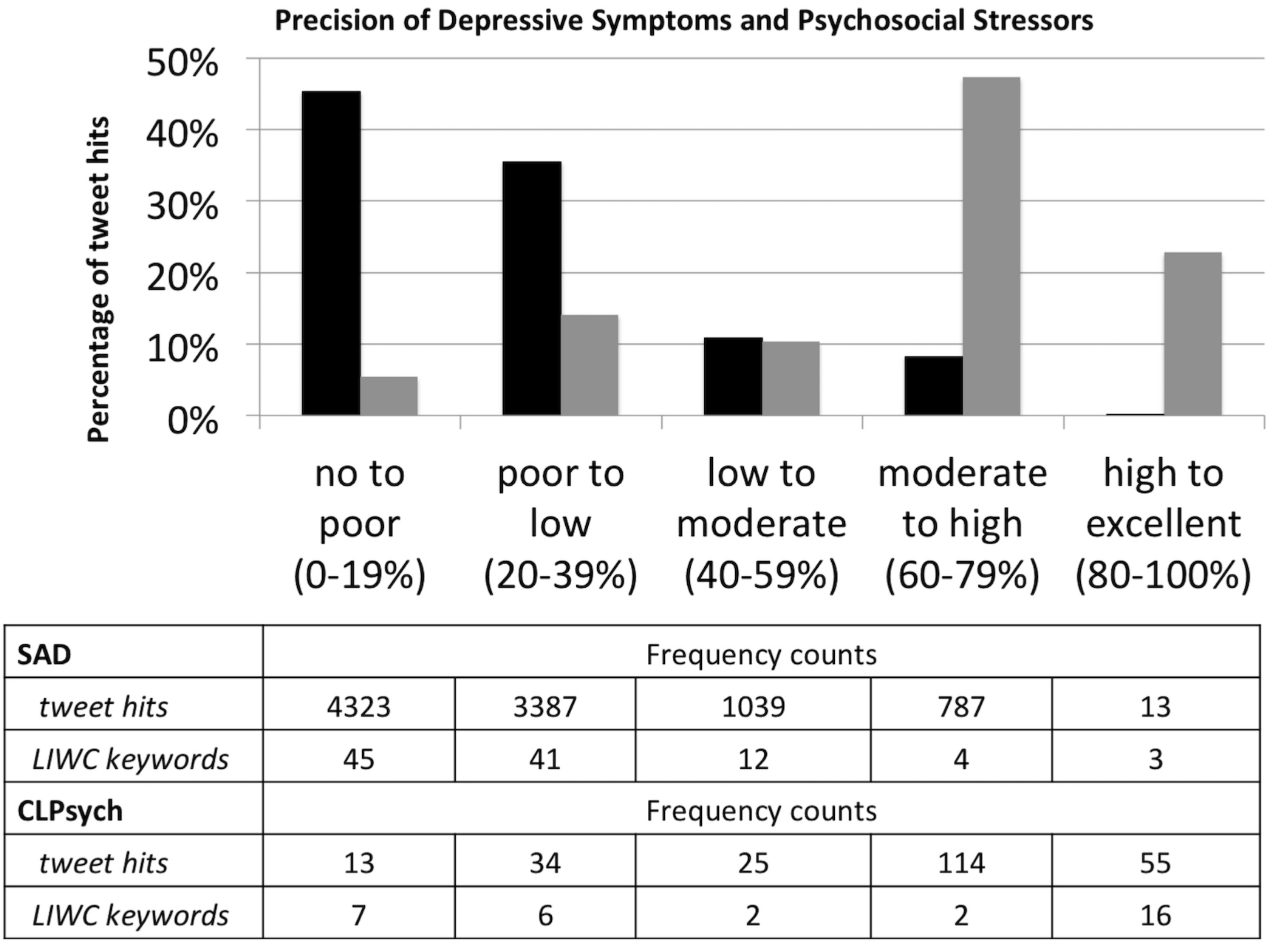
Table 3. For the SAD corpus, interannotator agreement (*F* scores) between annotators according to depressive symptoms and psychosocial stressors. — means category not observed by annotators.

Depression categories	A1/A2, (%)	A2/A3, (%)	A1/A3, (%)
Overall	81	78	76
No evidence of depression	89	86	87
Symptoms			
Depressed mood	38	60	48
Anhedonia	–	–	–
Weight change or change in appetite	–	0	100
Disturbed sleep	100	50	0
Psychomotor agitation or retardation	–	–	–
Fatigue or loss of energy	74	78	94
Feelings of worthlessness or excessive inappropriate guilt	0	29	68
Diminished ability to think or concentrate, indecisiveness	100	–	0
Recurrent thoughts of death, suicidal ideation	100	100	75
Stressors			
Problems with expected life course with respect to self	0	0	0
Problems with primary support group	0	40	36
Problems related to the social environment	23	42	58
Educational problems	–	50	0
Occupational problems	–	0	–
Housing problems	–	0	–
Economic problems	–	67	50
Problems with access to health care	–	0	–
Problems related to the legal system and crime	–	0	0
Other psychosocial and environmental problems	–	0	0
Weather	–	100	–
Media	50	0	67

Learning the Predictive Value of Depression-Related Keywords

For the SAD corpus, of the 110 unique depression-related keywords, 105 keywords were found corresponding to 9549 nonmutually exclusive tweet hits. We observed a range of precision across depression-related keyword hits: 45.27% (4323/9549) zero to poor, 35.47% (3387/9549) poor to low, 10.88% (1039/9549) low to moderate, 8.24% (787/9549) moderate to high, and 0.14% (13/9549) high to excellent (Figure 3). For the CLPsych corpus, the 35 unique depression-related keywords found correspond to 241 nonmutually exclusive tweet hits. We observed a range of precision across depression-related keyword hits: 5.40% (13/241) zero to poor, 14.11% (34/241) poor to low, 10.37% (25/241) low to moderate, 47.30% (114/241) moderate to high, and 22.82% (55/241) high to excellent.

Figure 3. Distribution of tweet hits by precision with LIWC Keyword counts for each corpus. Black bars=SAD corpus; Gray bars= CLPsych corpus. SAD: Depressive Symptom and Psychosocial Stressors Acquired Depression.

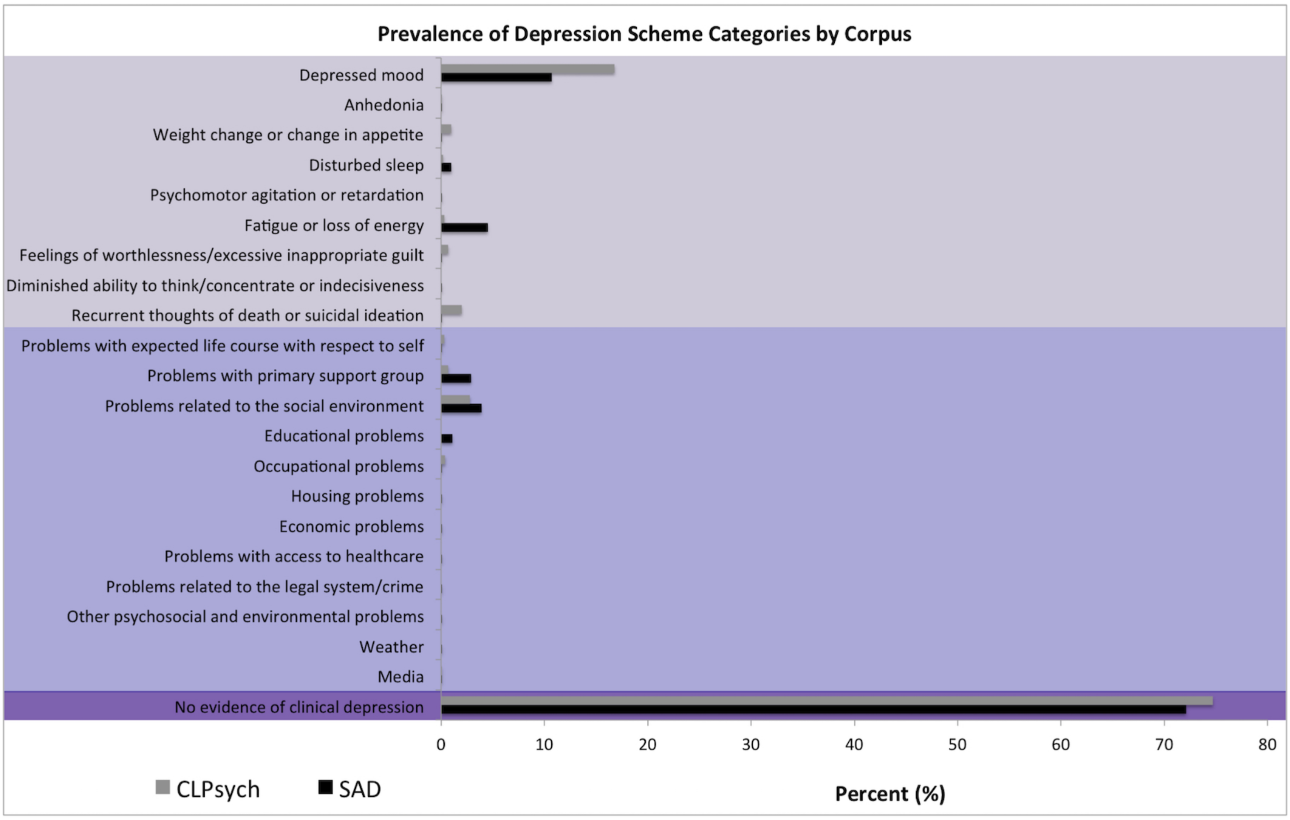


SAD	Frequency counts				
tweet hits	4323	3387	1039	787	13
LIWC keywords	45	41	12	4	3
CLPsych	Frequency counts				
tweet hits	13	34	25	114	55
LIWC keywords	7	6	2	2	16

Exploring the Frequency of Depressive Symptoms and Psychosocial Stressors

The SAD corpus consists of 9300 tweets. Of these tweets, 9293 were annotated with one or more categories from our scheme: 1 category (98.11%, 9117/9293), 2 categories (1.86%, 173/9293), and 3 or more categories (<1%, 3/9293). Overall, we observed a total of 9473 category annotations with the following distribution of categories annotated per tweets. A total of 72.09% (6829/9473) of annotations represent *no evidence of clinical depression* (Figure 4). Of the 27.91% (2644/9473) annotations that contained *evidence of clinical depression*, 18.20% (1724/9473) represented depressive symptoms and 9.71% (920/9473) represented psychosocial stressors. The CLPsych corpus consists of 1019 tweets. All tweets were annotated with only 1 category from our scheme. About 74.68% (761/1019) of annotations represent *no evidence of clinical depression*. Of the 25.32% (258/1019) annotations that contained *evidence of clinical depression*, 19.04% (194/1019) represented depressive symptoms and 6.28% (64/1019) represented psychosocial stressors.

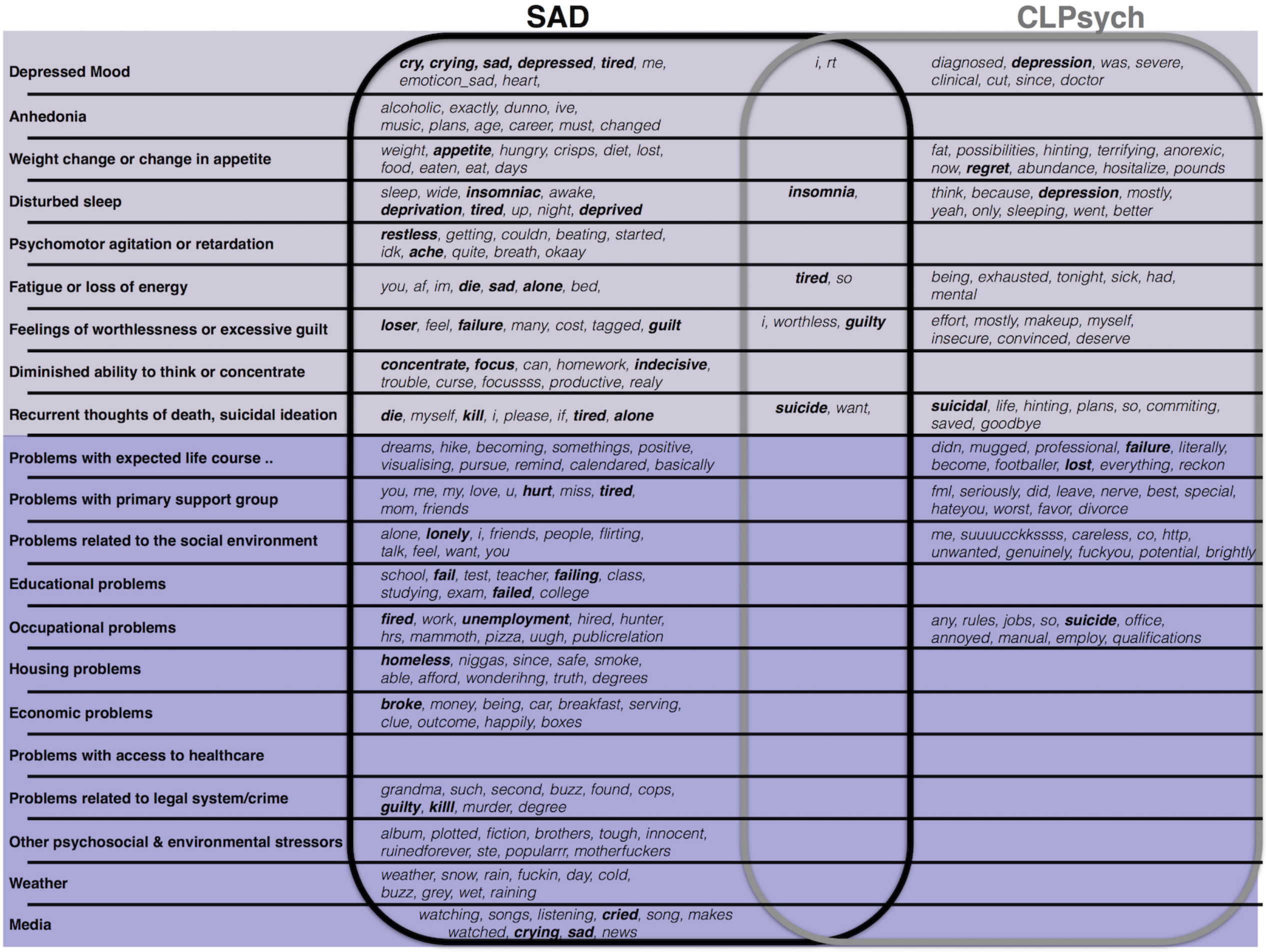
Figure 4. Prevalence of categories by corpus. Light purple: depressive symptoms, medium purple: psychosocial stressors, dark purple: no evidence of clinical depression.



Determining Predictive Word Features for Depressive Symptoms and Psychosocial Stressors

For the SAD corpus, 31 words were identified as the most informative features for classifying tweets for 11 depressive symptoms and psychosocial stressor categories (Figure 5). About 19 of these terms are also covered by the original LIWC keyword list.

Figure 5. Most informative terms classified with associated depressive symptoms and psychosocial stressors. Shared terms occur at the intersect of the circled lists.

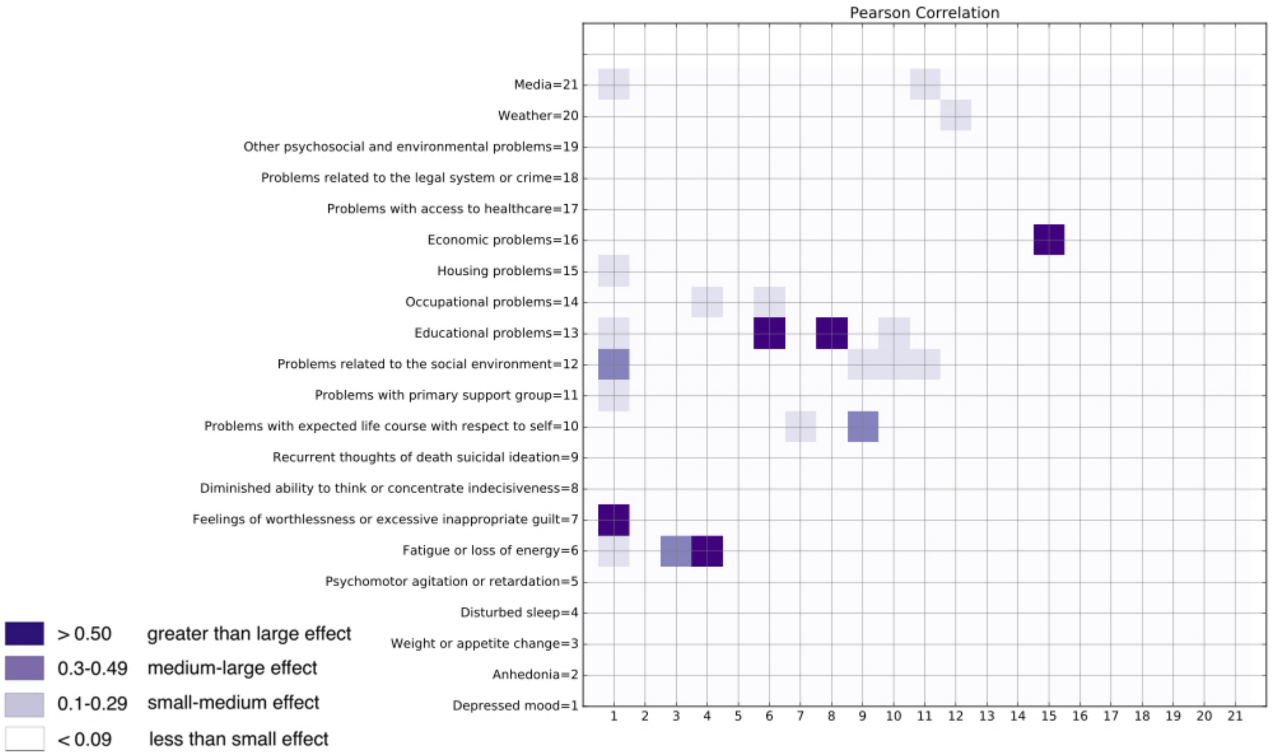


	SAD		CLPsych
Depressed Mood	**cry, crying, sad, depressed**, **tired**, me, emoticon_sad, heart,	i, rt	diagnosed, **depression**, was, severe, clinical, cut, since, doctor
Anhedonia	alcoholic, exactly, dunno, ive, music, plans, age, career, must, changed		
Weight change or change in appetite	weight, **appetite**, hungry, crisps, diet, lost, food, eaten, eat, days		fat, possibilities, hinting, terrifying, anorexic, now, **regret**, abundance, hositalize, pounds
Disturbed sleep	sleep, wide, **insomniac**, awake, **deprivation**, **tired**, up, night, **deprived**	**insomnia**,	think, because, **depression**, mostly, yeah, only, sleeping, went, better
Psychomotor agitation or retardation	**restless**, getting, couldn, beating, started, idk, **ache**, quite, breath, okaay		
Fatigue or loss of energy	you, af, im, **die**, **sad**, **alone**, bed,	**tired**, so	being, exhausted, tonight, sick, had, mental
Feelings of worthlessness or excessive guilt	**loser**, feel, **failure**, many, cost, tagged, **guilt**	i, worthless, **guilty**	effort, mostly, makeup, myself, insecure, convinced, deserve
Diminished ability to think or concentrate	**concentrate, focus**, can, homework, **indecisive**, trouble, curse, focussss, productive, realy		
Recurrent thoughts of death, suicidal ideation	**die**, myself, **kill**, i, please, if, **tired**, **alone**	**suicide**, want,	**suicidal**, life, hinting, plans, so, commiting, saved, goodbye
Problems with expected life course ..	dreams, hike, becoming, somethings, positive, visualising, pursue, remind, calendared, basically		didn, mugged, professional, **failure**, literally, become, footballer, **lost**, everything, reckon
Problems with primary support group	you, me, my, love, u, **hurt**, miss, **tired**, mom, friends		fml, seriously, did, leave, nerve, best, special, hateyou, worst, favor, divorce
Problems related to the social environment	alone, **lonely**, i, friends, people, flirting, talk, feel, want, you		me, suuuucckkssss, careless, co, http, unwanted, genuinely, fuckyou, potential, brightly
Educational problems	school, **fail**, test, teacher, **failing**, class, studying, exam, **failed**, college		
Occupational problems	**fired**, work, **unemployment**, hired, hunter, hrs, mammoth, pizza, uugh, publicrelation		any, rules, jobs, so, **suicide**, office, annoyed, manual, employ, qualifications
Housing problems	**homeless**, niggas, since, safe, smoke, able, afford, wonderihng, truth, degrees		
Economic problems	**broke**, money, being, car, breakfast, serving, clue, outcome, happily, boxes		
Problems with access to healthcare			
Problems related to legal system/crime	grandma, such, second, buzz, found, cops, **guilty**, **killl**, murder, degree		
Other psychosocial & environmental stressors	album, plotted, fiction, brothers, tough, innocent, ruinedforever, ste, popularrr, motherfuckers		
Weather	weather, snow, rain, fuckin, day, cold, buzz, grey, wet, raining		
Media	watching, songs, listening, **cried**, song, makes watched, **crying**, **sad**, news		

Assessing Correlations Between Depressive Symptoms and Psychosocial Stressors

In terms of depressive symptoms and psychosocial stressors, we observed 5 pairs with higher than large correlations, 3 pairs with medium to large correlations, and 13 with small correlations (Figure 6). Furthermore, all other possible combinations were either of low effect (≤0.09) or not observed in the corpus. Specifically, *fatigue or loss of energy* demonstrated large effect with *disturbed sleep* and *educational problems*. *Depressed mood* had large effect with *feelings of worthlessness* or *excessive inappropriate guilt*. *Educational problems* had large effect with *fatigue or loss of energy* and *diminished ability to think or concentrate* and *indecisiveness*. *Housing problems* and *economic problems* also demonstrated a large effect.

Figure 6. SAD heat map of tweet-level, depressive symptom, and psychosocial stressor cooccurrences. Darker means larger measure of Cohen effect size; lighter means smaller measure of Cohen effect size. The number that indexes the category on the y-axis also corresponds to the category for the x-axis. For example, if "Depressed mood=1" appears on the y-axis, then "1" on the x-axis corresponds to the category "Depressed mood." SAD: Depressive Symptom and Psychosocial Stressors Acquired Depression.



Discussion

Principal Findings

In summary, several depressive symptoms and psychosocial stressor categories could be observed in the corpus. For tweets containing two or more categories, we found large correlations between some depressive symptoms and psychosocial stressor categories. Our assessment also suggests that keyword queries alone might not be suitable for public health monitoring.

Characterizing the Corpus

We conducted an annotation study to investigate methods for effective data collection and understand how people tweet about depression on Twitter. We observed similar average number of tokens and standard deviations for both the SAD and CLPsych corpora (Table 2).

Validating the Annotation Scheme

In order to address these aims, we applied our scheme to the SAD corpus. We observed that annotators are able to discern tweets representing *no evidence of clinical depression* and achieve high overall *F* scores (acceptable within the NLP community [46]; Table 3). However, we observed variable *F* scores for depressive symptoms and psychosocial stressors, which we attribute to the lower prevalence of these categories in the corpus.

Learning the Predictive Value of Depression-Related Keywords

Specifically, we assessed the predictive value of depression-related keywords for effective data collection because the mechanism for collecting data, the Twitter API, can only apply keywords to retrieve relevant tweets. We observed different distributions of precision between the SAD and CLPsych corpora (Figure 3). For the SAD corpus, most depression-related keywords demonstrated zero to poor to low precision. In contrast, the CLPsych corpus, most depression-related keywords demonstrated moderate to high to excellent precision. We hypothesize that the depression-related keywords have better precision because of the lack of ambiguity in their usage due to contextual grounding with the self-reported diagnosis ("I was diagnosed with depression"). Specifically, for the SAD corpus, less than 1% of the tweets were classified as high to excellent precision that were identified by querying tweets with 3 depression-related keywords: "inferior," "dishearten," and "restless." For example, "Everyday leaves me feeling more hopeless and restless." In contrast, for the CLPsych corpus, more than 22% of the tweets were classified with high to excellent precision which were identified by querying tweets with 15 depression-related keywords such as "inferior," "dishearten," "depressants," "suicidal," "tired," "miserable," "depressive," "suicide," "divorce," "unhappy," "heartbreak," "lonely," "insomnia," "depressing," and "hurts." For example, "I always feel insecure and inferior to everyone in my life." From this assessment, we will leverage these depression-related keywords to query tweets related to depressive symptoms: *depressed mood*, *disturbed sleep*, *fatigue or loss of energy*, *feelings of worthlessness or excessive inappropriate guilt*, as well as psychosocial stressors: *recurrent thoughts of death*, *suicidal ideation*, *problems with primary support group*, and *problems related to the social environment*.

Exploring the Frequency of Symptoms and Psychosocial Stressors

Overall, we observed similar distributions of *no evidence of clinical depression* and *evidence of clinical depression* categories as well as depressive symptoms and psychosocial stressors between the SAD and CLPsych Corpora (Figure 4). We observed a skewed distribution of depressive symptoms and psychosocial stressors categories in both corpora. The most prevalent category for both corpora was *no evidence of clinical depression* meaning for every 10 tweets reviewed 7 were not relevant. This finding suggests that our a priori depression-related keyword lexicon was insufficient for consistently identifying depression-related tweets and that natural language processing methods will be required to increase accuracy.

When evidence of clinical depression was identified for both the SAD and CLPsych corpora, tweets more often described depressive symptoms rather than psychosocial stressors. This finding suggests that Twitter users may be more comfortable or feel an immediate need to describe their current mental state and physical feelings (eg, “I can’t concentrate”) rather than the psychosocial stressors that may have given rise to these depressive symptoms (eg, “I can’t concentrate because of my recent car accident”). In terms of depressive symptoms, both corpora contained *depressed mood* as the most prevalent depressive symptom. However, for the SAD corpus, the following second and third most prevalent depressive symptoms included *fatigue or loss of energy* and *disturbed sleep*; in contrast to the CLPsych corpus, in which the following second and third most prevalent depressive symptoms included *weight change or change in appetite* and *feelings of worthlessness or excessive inappropriate guilt*. In terms of psychosocial stressors, both corpora contained *problems related to the social environment* and *problems with primary support group*. However, for the SAD corpus, the third most prevalent psychosocial stressor included *educational problems*; whereas, for the CLPsych corpus, the third most prevalent psychosocial stressors included *recurrent thoughts of death and suicidal ideation*. The SAD depressive symptoms and psychosocial stressor distributions are unsurprising and mirror the distributions found in our pilot annotation effort [39,47].

Determining Predictive Word Features for Depressive Symptoms and Psychosocial Stressors

To expand on our data acquisition approach and supplement the depression-related keyword lexicon, we also conducted a feature selection study to identify words most characteristic of each depression symptom and psychosocial stressor with the aim of identifying new keywords not already present in our lexicon of depression-related keywords. For the SAD corpus, only one category— *problems with access to health care* —had too few mentions to learn new keywords (Figure 5). Of the most informative keywords identified, most were absent from our handcrafted depression-keyword lexicon, suggesting that some new words could be useful for pulling relevant tweets for most depressive symptoms and psychosocial stressor categories. For the CLPsych corpus, we observed many new informative words. However, only about half of the categories had more than 2 mentions. Few depression-related words were shared between the SAD and CLPsych corpora, suggesting that we may still learn new words. Similar to Coppersmith et al [32] and Preotuic-Pietro et al [34], our work indicates that greater use of personal pronouns could indicate an increased focus on the self. We also observed words for many depressive symptoms and psychosocial stressors associated with *anxiety* and *anger* and biological states such as *health* and *death*. These new words are promising; however, we leave it to future studies to determine their precision or recall on a new, unseen Twitter dataset.

Assessing Correlations Between Depressive Symptoms and Psychosocial Stressors

In terms of depressive symptoms and psychosocial stressors, we observed 5 pairs with higher than large effects (Figure 6). Specifically, *fatigue or loss of energy* demonstrated large effects with another depressive symptom of *disturbed sleep* and psychosocial stressor of *educational problems*. Our analysis suggests that individuals expressing chronic fatigue describe this symptom affecting their quality of life including difficulties in managing sleep and nutrition, productivity at work or school, and interactions with others [20]. *Depressed mood* demonstrated large effect with another depressive symptom of *feelings of worthlessness or excessive inappropriate guilt*. Other interesting and intuitive findings are that *educational problems* exhibited large effects with other symptoms of *fatigue or loss of energy* and *diminished ability to think or concentrate and indecisiveness*, suggesting that if an individual experiences problems during his or her academic studies it could be attributed to tiredness and the inability to concentrate on subject matter. *Housing problems* and *economic problems* also demonstrated large effect, a fact that makes sense intuitively if we consider that an individual experiencing economic problems may encounter difficulties maintaining a home.

Limitations

For the SAD corpus, we cannot confirm whether an individual Twitter user has or has not received a formal diagnosis of depression. However, many individuals go undiagnosed for depression; therefore, one advantage of this methodology is that it could capture relevant symptomology without a formal diagnosis. However, it is important to be clear that for ethical reasons (eg, individual privacy) the intent of this tool is not to diagnose depression or attempt to intervene at the individual level, but rather to estimate and report the prevalence of depression symptoms at the population level over time in the United States. Furthermore, the correlational analysis performed on the SAD corpus could not be performed for the smaller CLPsych corpus, as we did not observe more than one depression symptom or psychosocial stressor associated with each tweet.

Comparison With Prior Work

Since our pilot study on a dataset of 500 depression-related tweets [39,47], little research has been conducted specifically to qualitatively (rather than computationally) understand the range of depression-related symptoms that manifested in Twitter data. An important exception is Cavazos-Rehg et al, who used a qualitative technique to study 2000 randomly selected tweets containing one or more depression-related keywords (depressed, #depressed, depression, #depression), finding that two-thirds of the tweets either described depressive symptoms, or expressed thoughts consistent with depression [48]. This study complements and builds on that reported in Cavazos-Rehg et al in several key ways. First, the primary dataset leveraged in this study is almost 5 times larger than that used by Cavazos-Rehg et al (9300 tweets and 2000 tweets, respectively). Second, the dataset used in this study was created using a variety of keywords related to depression and depressive symptoms (110 in total) rather than Cavazos-Rehg et al's use of lexical variants of the word "depression." Third, this study extends beyond the analysis of DSM 5 depressive symptoms to include psychosocial stressors derived from DSM-IV Axis IV [4] (eg, *educational problems*, *occupational problems*, *problems related to the social environment*). Finally, this study is designed to investigate correlations between depression symptoms and psychosocial stressors.

This study has 2 main goals: First, to provide insights into how users express depressive symptoms on Twitter; and second, to create a dataset (ie, an annotated corpus of depression-related tweets) suitable for both training and testing natural language processing algorithms to automate the process of identifying tweets manifesting evidence of depression symptoms. Although the dataset will not be openly available, the resulting, trained and tested natural language processing symptom classifiers will be openly available in the near future. These classifiers may be used to estimate and report the prevalence of other mental health disorders (eg, anxiety and eating disorders) by encoding shared symptoms and stressors leveraging similar language patterns from social media [33].

Conclusions

We conducted a large-scale annotation study to investigate methods for effective data collection and understand how people tweet about depression on Twitter with the twin goals of (1) providing insights into how users express depressive symptoms on Twitter and (2) creating a dataset (ie, an annotated corpus of depression-related tweets) suitable for both training and testing natural language processing algorithms to automate the process of identifying tweets manifesting evidence of depression symptoms. We successfully developed an annotation scheme and an annotated corpus, the SAD corpus, consisting of 9300 tweets randomly selected from the Twitter API using depression-related keywords. Although the majority of tweets containing relevant keywords were nonindicative of depressive symptoms, several depressive symptoms and psychosocial stressor categories were observed including *depressed mood* and *fatigue or loss of energy*. In tweets containing two or more categories, we found correlations between some depressive symptoms and psychosocial stressor categories.

In summary, our analyses suggest that keyword queries alone might not be suitable for public health monitoring because the context can change the meaning of a keyword in a statement. However, postprocessing approaches could be useful for reducing the noise and improving the signal needed to detect depression symptoms using social media. We are actively investigating machine-learning based postprocessing as an approach to improve the precision of detecting depressive symptoms and psychosocial stressors [49,50].

Acknowledgments

Research reported in this publication was supported by the National Library of Medicine of the (United States) National Institutes of Health under award numbers K99LM011393 and R00LM011393. All authors reviewed, edited, and approved the resulting manuscript. This study was granted an exemption from review by the University of Utah Institutional Review Board (IRB 00076188). Note that in order to protect tweeter anonymity, we have not reproduced tweets verbatim. Example tweets shown were generated by the researchers as exemplars only. We thank Dr Murray Stein (University of California San Diego, Department of Psychiatry) and Dr Gary Tedeschi (California Smokers Helpline) for their comments on an early draft of the annotation scheme described in this paper. Finally, we thank our anonymous reviewers for their invaluable feedback.

Authors' Contributions

DM, CB, and MC developed the schema. DM and MC designed the study with GS providing statistical support. HS, TC, and DM annotated the SAD corpus. GC led the annotation of the CLPsych corpus. DM completed the corpus analysis. DM and MC led the writing of the manuscript.

Conflicts of Interest

GC is the founder and chief executive officer of the company, Qntfy. Qntfy provided support in the form of salary for the author GC, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Library of Medicine or the National Institutes of Health. There are no patents, products in development or marketed.

References

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). J Am Med Assoc 2003 Jun 18;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095] [Medline: 12813115]
2. Murray CJ, Atkinson C, Bhalla K, Birbeck G, Burstein R, Chou D, Foreman, Lopez, Murray, Dahodwala, Jarlais, Fahami, Murray, Jarlais, Foreman, Lopez, Murray, US Burden of Disease Collaborators. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. J Am Med Assoc 2013 Aug 14;310(6):591-608. [doi: 10.1001/jama.2013.13805] [Medline: 23842577]
3. Diagnostic and Statistical Manual of Mental Disorders DSM-5, Fifth Edition. Arlington, Va: American Psychiatric Publishing; May 18, 2013:1-947.
4. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), 4th Edition. Arlington, Va: American Psychiatric Association; 2000:1-886.
5. Dredze M. How social media will change public health. IEEE Intell Syst 2012;27(4):81-84. [doi: 10.1109/MIS.2012.76]
6. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
7. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
8. Paul MJ, Dredze M, Broniatowski D. PLoS Curr. 2014. Twitter improves influenza forecasting URL: http://currents.plos.org/outbreaks/article/twitter-improves-influenza-forecasting/ [accessed 2017-02-12] [WebCite Cache ID 6oELxhbRi]
9. Myslín M, Zhu S, Chapman W, Conway M. Using twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
10. Hanson CL, Cannon B, Burton S, Giraud-Carrier C. An exploration of social circles and prescription drug abuse through Twitter. J Med Internet Res 2013;15(9):e189 [FREE Full text] [doi: 10.2196/jmir.2741] [Medline: 24014109]
11. Chen AT, Zhu S, Conway M. What online communities can tell us about electronic cigarettes and hookah use: a study using text mining and visualization techniques. J Med Internet Res 2015;17(9):e220 [FREE Full text] [doi: 10.2196/jmir.4517] [Medline: 26420469]
12. Conway M, O'Connor D. Social media, big data, and mental health: current advances and ethical implications. Curr Opin Psychol 2016 Jun;9:77-82 [FREE Full text] [doi: 10.1016/j.copsyc.2016.01.004] [Medline: 27042689]
13. Pew Research Center. Mobile Messaging and Social Media 2015 URL: http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/ [accessed 2016-12-19] [WebCite Cache ID 6msTpi0ri]
14. GNIP. URL: https://gnip.com/ [accessed 2016-12-19] [WebCite Cache ID 6msTug41S]
15. Ritter A, Clark S, Mausam, Etzioni O. Named entity recognition in tweets: an experimental study. 2011 Jul 27 Presented at: Proceedings of the Conference on Empirical Methods on Natural Language Processing; July 27-31, 2011; Edinburgh, United Kingdom p. 1524-1534.
16. Mikal J, Hurst S, Conway M. Ethical issues in using Twitter for population-level depression monitoring: a qualitative study. BMC Med Ethics 2016 Apr 14;17:22 [FREE Full text] [doi: 10.1186/s12910-016-0105-5] [Medline: 27080238]
17. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006 Nov;3(11):e442 [FREE Full text] [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]
18. CDC. Behavioral risk factor surveillance system survey data; 2008, 2009, 2012 URL: https://www.cdc.gov/brfss/ [accessed 2016-12-19] [WebCite Cache ID 6msTxdJyk]
19. Bloom D, Cafiero E, Jae-Llopis E, Abrahams-Gessel S, Bloom L, Fathima S, et al. Hsph.harvard. 2012. The global economic burden of noncommunicable diseases URL: https://www.hsph.harvard.edu/program-on-the-global-demography-of-aging/WorkingPapers/2012/PGDA_WP_87.pdf [accessed 2016-12-19] [WebCite Cache ID 6msU396DV]
20. NIMH. NIH. Mental health information: depression URL: https://www.nimh.nih.gov/health/topics/depression/index.shtml [accessed 2016-12-19] [WebCite Cache ID 6msU5WgpI]
21. Sanderson CL, Feng S, Cañar J, McGlinchey FM, Tercyak KP. Social and behavioral correlates of cigarette smoking among mid-Atlantic Latino primary care patients. Cancer Epidemiol Biomarkers Prev 2005 Aug;14(8):1976-1980 [FREE Full text] [doi: 10.1158/1055-9965.EPI-05-0141] [Medline: 16103447]
22. Witkiewitz K, Villarroel NA. Dynamic association between negative affect and alcohol lapses following alcohol treatment. J Consult Clin Psychol 2009 Aug;77(4):633-644 [FREE Full text] [doi: 10.1037/a0015647] [Medline: 19634957]
23. ten Hacken NH. Physical inactivity and obesity: relation to asthma and chronic obstructive pulmonary disease? Proc Am Thorac Soc 2009 Dec;6(8):663-667. [doi: 10.1513/pats.200907-070DP] [Medline: 20008872]
24. Coulombe JA, Reid GJ, Boyle MH, Racine Y. Sleep problems, tiredness, and psychological symptoms among healthy adolescents. J Pediatr Psychol 2011 Jan;36(1):25-35 [FREE Full text] [doi: 10.1093/jpepsy/jsq028] [Medline: 20421201]
25. SAMHSA. 2016. About population data/NSDUH URL: http://www.samhsa.gov/data/population-data-nsduh/about [accessed 2016-12-19] [WebCite Cache ID 6msU9SO3x]
26. CDC. Youth risk behavior surveillance system URL: https://www.cdc.gov/healthyyouth/data/yrbs/index.htm [accessed 2016-12-19] [WebCite Cache ID 6msUBVcbO]

27. CDC. BRFSS - Anxiety and Depression Optional Module URL: https://www.cdc.gov/mentalhealth/data_stats/pdf/BRFSS_Anxiety_and_Depression_Optional_Module.pdf [accessed 2016-12-19] [WebCite Cache ID 6msUFRuLW]
28. De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. : Association for Computing Machinery; 2013 May 02 Presented at: Proceedings of the 5th Annual ACM Web Science Conference; May 2-4, 2013; Paris, France p. 47-56.
29. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. 2013 Jun 08 Presented at: Proceedings of the 7th International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media; June 8-10, 2013; Boston, Massachusetts.
30. Schwartz HA, Eichstaedt J, Kern ML, Park G, Sap M, Stillwell D, et al. Towards assessing changes in degree of depression through Facebook. 2014 Jun 27 Presented at: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 27, 2014; Baltimore, Maryland p. 118-125.
31. Tausczik Y, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 2010;29(1):24-54. [doi: 10.1177/0261927X09351676]
32. Coppersmith G, Dredze M, Harman C. Quantifying mental health signals in Twitter. 2014 Jun 27 Presented at: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 27, 2014; Baltimore, Maryland p. 51-60.
33. Coppersmith G, Dredze M, Harman C, Hollingshead K. From ADHD to SAD: analyzing the language of mental health on Twitter through self-reported diagnoses. 2015 Jun 05 Presented at: Proceeding of 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, Colorado p. 1-8.
34. Preotiuc-Pietro D, Eichstaedt J, Park G, Sap M, Smith L, Tobolsky V, et al. The role of personality, age, and gender in tweeting about mental illness. 2015 Jun 05 Presented at: Proceeding of 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, Colorado p. 21-23.
35. Baer L, Jacobs D, Meszler-Reizes J, Blais M, Fava M, Kessler R, et al. Development of a brief screening instrument: the HANDS. Psychother Psychosom 2000;69(1):41. [Medline: 10601833]
36. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [Medline: 11556941]
37. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 01;54(5):573-583. [Medline: 12946886]
38. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
39. Mowery D, Bryan C, Conway M. Toward developing an annotation scheme for depressive disorder symptoms: a preliminary study using Twitter data. In: Proceeding of 2nd Workshop on Computational Linguistics and Clinical Psychology. 2015 Jun 05 Presented at: From Linguistic Signal to Clinical Reality. Association for Computational Linguistics; June 5, 2015; Denver, Colorado p. 89-98.
40. Coppersmith G, Dredze M, Harman C, Hollingshead K, Mitchell M. CLPsych 2015 shared task: depression and PTSD on Twitter. In: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology.: Association for Computational Linguistics; 2015 Jun 05 Presented at: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, Colorado p. 31-39 URL: https://github.com/clpsych/shared_task
41. Hripcsak G, Rothschild AS. Agreement, the f-measure, and reliability in information retrieval. J Am Med Inform Assoc 2005;12(3):296-298 [FREE Full text] [doi: 10.1197/jamia.M1733] [Medline: 15684123]
42. Witten I, Frank E, Hall M. Data Mining: Practical Machine Learning Tools and Techniques. 3rd ed. Boston, Masschusetts: Morgan Kaufmann Publishers; 2011.
43. Conway M. Mining a corpus of biographical texts using keywords. Literary and Linguistic Computing 2010;25(1):23-35. [doi: 10.1093/llc/fqp035]
44. Yang Y, Pedersen J. A comparative study on feature selection in text categorization. : Morgan Kaufmann Publishers; 1997 Presented at: Proceedings of the Fourteenth Annual Conference on Machine Learning; 1997; Singapore, Singapore p. 412-420.
45. Cohen J. A power primer. Psychol Bull 1992;1(112):155-159. [Medline: 19565683]
46. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-174. [Medline: 843571]
47. Mowery D, Smith H, Cheney T, Bryan C, Conway M. Identifying depression-related tweets from Twitter public health monitoring. Open J Public Health Inform 2016;8(1):e144. [doi: 10.5210/ojphi.v8i1.6561]
48. 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 1;54:351-357. [doi: 10.1016/j.chb.2015.08.023] [Medline: 26392678]
49. Mowery D, Park A, Bryan C, Conway M. Towards automatically classifying depressive symptoms from Twitter data for population health. 2016 Dec 12 Presented at: Proceedings of the Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media; December 12, 2016; Osaka, Japan p. 182-191.

50. Mowery D, Bryan C, Conway M. Feature studies to inform the classification of depressive symptoms from Twitter data for population health. 2017 Feb 10 Presented at: Proceedings of the WSDM 2017 Workshop on Mining Online Health Reports; February 10, 2017; Cambridge, United Kingdom.

Abbreviations

API: Application Programming Interface
BRFSS: Behavioral Risk Factors Surveillance System
CLPsych: Computational Linguistics and Clinical Psychology
DSM-V: Diagnostic and Statistical Manual of Mental Disorders, Edition 5
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Edition IV
HANDS: Harvard Department of Psychiatry National Depression Screening Day Scale
LIWC: Linguistic Inquiry and Word Count
NLP: natural language processing
NSDUH: National Survey on Drug Use and Health
PHQ-5: Patient Health Questionnaire
QIDS-SR: Quick Inventory of Depressive Symptomatology
SAD: Depressive Symptom and Psychosocial Stressors Acquired Depression
YRBSS: Youth Risk Behavior Surveillance System

Edited by R Calvo; submitted 26.10.16; peer-reviewed by G Gkotsis, S Pagoto, M Paul; comments to author 05.12.16; accepted 24.01.17; published 28.02.17.

Please cite as:
Mowery D, Smith H, Cheney T, Stoddard G, Coppersmith G, Bryan C, Conway M
Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study
J Med Internet Res 2017;19(2):e48
URL: http://www.jmir.org/2017/2/e48/
doi:10.2196/jmir.6895
PMID:28246066

Original Paper

Experience With Direct-to-Patient Recruitment for Enrollment Into a Clinical Trial in a Rare Disease: A Web-Based Study

Jeffrey Krischer[1*], PhD; Peter F Cronholm[2*], MSCE, MD; Cristina Burroughs[1*], BS; Carol A McAlear[2*], MA; Renee Borchin[1*], BSW; Ebony Easley[2*], MPH; Trocon Davis[2*], BA; Joyce Kullman[3*], BS; Simon Carette[4*], MD; Nader Khalidi[5*], MD; Curry Koening[6*], MS, MD; Carol A Langford[7*], MHS, MD; Paul Monach[8*], MD, PhD; Larry Moreland[9*], MD; Christian Pagnoux[4*], MD; Ulrich Specks[10*], MD; Antoine G Sreih[2*], MD; Steven Ytterberg[10*], MD; Peter A Merkel[2*], MPH, MD; Vasculitis Clinical Research Consortium[11*]

[1]Rare Diseases Clinical Research Network Data Coordinating Center, Health Informatics Institute, University of South Florida, Tampa, FL, United States
[2]University of Pennsylvania, Philadelphia, PA, United States
[3]Vasculitis Foundation, Kansas City, MO, United States
[4]Mount Sinai Hospital, Toronto, ON, Canada
[5]St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada
[6]University of Utah, Salt Lake City, UT, United States
[7]Cleveland Clinic Foundation, Cleveland, OH, United States
[8]Boston University School of Medicine, Boston, MA, United States
[9]University of Pittsburgh, Pittsburgh, PA, United States
[10]Mayo Clinic, Rochester, MN, United States
[11]Vasculitis Clinical Research Consortium, Philadelphia, PA, United States
[*]all authors contributed equally

Corresponding Author:
Jeffrey Krischer, PhD
Rare Diseases Clinical Research Network Data Coordinating Center
Health Informatics Institute
University of South Florida
3650 Spectrum Blvd, Suite 100
Tampa, FL, 33612
United States
Phone: 1 813 396 9512
Fax: 1 877 775 7987
Email: Jeffrey.Krischer@epi.usf.edu

Abstract

Background: The target sample size for clinical trials often necessitates a multicenter (center of excellence, CoE) approach with associated added complexity, cost, and regulatory requirements. Alternative recruitment strategies need to be tested against this standard model.

Objectives: The aim of our study was to test whether a Web-based direct recruitment approach (patient-centric, PC) using social marketing strategies provides a viable option to the CoE recruitment method.

Methods: PC recruitment and Web-based informed consent was compared with CoE recruitment for a randomized controlled trial (RCT) of continuing versus stopping low-dose prednisone for maintenance of remission of patients with granulomatosis with polyangiitis (GPA).

Results: The PC approach was not as successful as the CoE approach. Enrollment of those confirmed eligible by their physician was 10 of 13 (77%) and 49 of 51 (96%) in the PC and CoE arms, respectively (*P*=.05). The two approaches were not significantly different in terms of eligibility with 34% of potential participants in the CoE found to be ineligible as compared with 22% in the PC arm (*P*=.11) nor in provider acceptance, 22% versus 26% (*P*=.78). There was no difference in the understanding of the trial as reflected in the knowledge surveys of individuals in the PC and CoE arms.

Conclusions: PC recruitment was substantially less successful than that achieved by the CoE approach. However, the PC approach was good at confirming eligibility and was as acceptable to providers and as understandable to patients as the CoE approach. The PC approach should be evaluated in other clinical settings to get a better sense of its potential.

(J Med Internet Res 2017;19(2):e50) doi:10.2196/jmir.6798

KEYWORDS

clinical trial; research subject recruitment; social media; direct-to-consumer advertising; granulomatosis with polyangiitis

Introduction

Despite the emphasis given to randomized controlled trials (RCTs) as the gold standard for the evaluation of new and promising therapies, it is well recognized that sufficient numbers of potential study participants are usually not available at a single institution, necessitating the organization of multicenter studies, development of specialized infrastructure (eg, study staff, training of site personnel, data transfer practices), layers of additional administrative work (eg, multi-institutional agreements and subcontracts, applications to multiple institutional review boards [IRBs], investigator meetings), and all the effort required to maintain study team cohesiveness and momentum across multiple sites. Yet, even with such high levels of investment, accrual into RCTs often fails to meet enrollment goals, even after enrollment periods are markedly extended. As many as 40% of all trials and 71% of phase 3 trials supported by the National Cancer Institute’s clinical trials program fail to ever achieve their target accrual [1,2]. Concern over the failure to achieve targeted accrual numbers extends to all trials nationally, with 19% of RCTs closing without having achieved at least 85% of their target accrual [3]. Among reasons significantly associated with unsuccessful accrual are increased number of eligibility requirements (presumably limiting the number of eligible individuals), less number of research sites, nonindustry funding, and a nonplacebo comparison arm.

Expanding the number of treatment sites in a traditional center of excellence (CoE) model for conducting RCTs is associated with significant costs. Even if multiple sites can be coordinated, a major problem with the traditional CoE model is that the majority of potential study participants are still out of reach of the few geographically-limited clinical centers involved in the trial. This is especially the case in trials targeting individuals with rare diseases [4]. Catchment areas and referral patterns are often related to a lack of experts in the disease under study, institutional competition, insurance agreements, or personal referral networking among physicians. Novel ideas to reach a larger and perhaps less preselected population could reduce the time and cost of clinical trials, increase the generalizability and social value of their findings, and increase the likelihood that important clinical questions could be addressed more quickly and successfully to better support advancing new beneficial therapies to affected individuals. Even more importantly, such methods would revolutionize clinical research in rare diseases.

These factors motivate the consideration of novel methods to recruit potentially eligible individuals to RCTs. The pharmaceutical industry makes extensive use of direct-to-consumer advertising of approved prescription drugs [5,6] with a presumed objective to attract patients to their drugs, as well as to inform about treatment alternatives and encourage communication with health care providers. Although regulated by the Food and Drug Administration (FDA), this marketing strategy is not without its detractors, especially when marketing begins soon after approval and there remains a need to educate providers on the appropriate use of the drug, and more data on possible untoward side effects is needed [7].

Although investigators rarely have the budget to employ the same advertising strategies as does the pharmaceutical industry, the rapid growth of the Internet and social media provide inexpensive, and in some cases, free access to millions of potential research participants. The value of advertising over the Internet is a major business strategy and investigators have taken note to use this methodology to broaden their reach to potentially eligible individuals and simultaneously overcome the barriers to enrollment into RCT [8-14].

This paper describes an effort to test whether direct recruitment of study individuals using social marketing strategies (ie, Web-based tools, online patient communities, Facebook, Twitter, Google+, others) to reach a target population and provide an interactive Web-based method to engage, educate, enroll, and obtain informed consent, provides a viable option for recruitment and enrollment into an RCT. Notably, the approach also included elements designed to reduce known barriers to enrollment [15] by having only one IRB, reducing study burden by not requiring treating physicians to be involved in the research, and asking an important clinical question, with 2 equally attractive treatment arms that were considered accepted standards of care.

Methods

Rationale for The Assessment of Prednisone in Remission (TAPIR) Trial

To test direct patient recruitment (patient-centric, PC) and compare it with the traditional CoE recruitment model, a clinical trial was designed for use as the study setting. The Assessment of Prednisone in Remission (TAPIR) trial tests whether patients with granulomatosis with polyangiitis (GPA; Wegener’s) have better outcomes after their GPA is well-controlled (in remission) if they stay on a maintenance dose of 5 mg/day of prednisone or fully come off of prednisone (0 mg/day; Clinicaltrials.gov NCT01940094 and NCT01933724).

Studies in the last 20 years have addressed the use of immunosuppressive medications in GPA. Unlike immunosuppressive medications, the use of prednisone has not been rigorously evaluated. There is little evidence to guide the use of prednisone and there is considerable practice pattern variation, especially after the induction of remission. Of particular debate is whether low-dose prednisone contributes to maintaining the remission of GPA. Some experts support the use of long-term, low-dose prednisone, claiming improved disease control, a subsequent reduction in the exposure to toxic immunosuppressive medications, fewer periods of exposure to high-dose prednisone, and a reduction in the accumulation of disease-related scarring. Others argue that the use of long-term, low-dose prednisone is ineffective at reducing relapses and exposes patients to the potential toxicity of high cumulative doses of prednisone. The efficacy of long-term, low-dose prednisone for the treatment of GPA to prevent relapses or reduce treatment-related toxicity is a matter of continued debate [16].

Trial Setting

The setting for this RCT is the Vasculitis Clinical Research Consortium (VCRC) **,** a founding member of the Rare Diseases Clinical Research Network (RDCRN) [17] and the major clinical research infrastructure in North America for the study of vasculitis. The work conducted by the VCRC includes clinical trials, outcome measures development, large cohort and clinical epidemiologic studies, biospecimen collection and repository, translational investigations including biomarker discovery and genomics, and research training. The US and Canada VCRC vasculitis clinical centers (ie, CoE sites) participated in this study. Simultaneously, the PC approach has been implemented by the VCRC and the RDCRN Data Management and Coordinating Center (DMCC) for cross-sectional studies and patient communication using Web-based tools and the Vasculitis Patient Contact Registry, which consists of nearly 3000 individuals in the United States with vasculitis, of whom 1458 report a diagnosis of GPA.

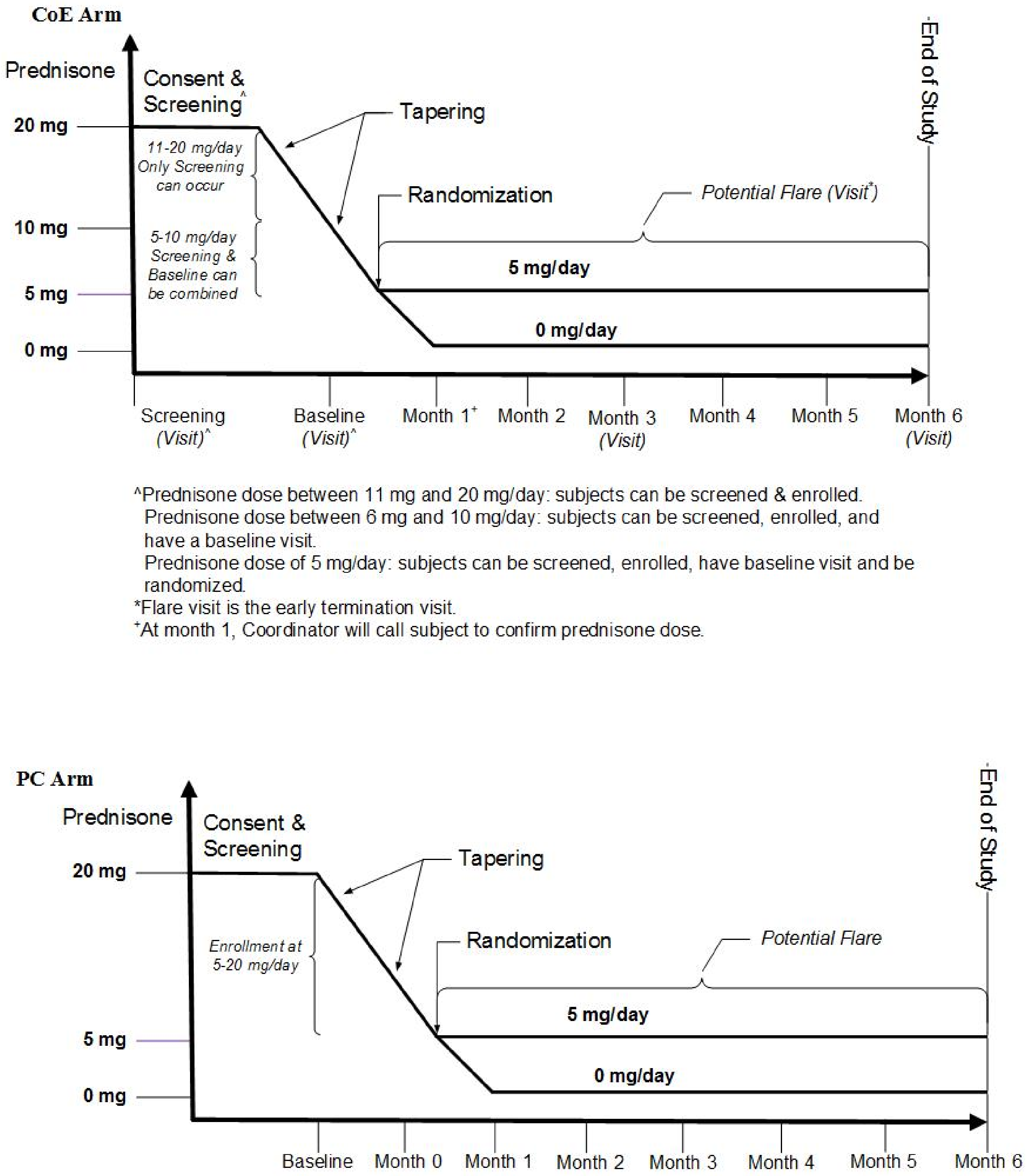
The TAPIR Protocol

To be eligible, patients must have an established diagnosis of GPA (verified by medical record review by protocol oversight management team) and meet at least two of the 5 modified American College of Rheumatology (ACR) criteria [18]. Patients must have had active disease within the prior 12 months (initial presentation or relapse) that at time of active disease required treatment with prednisone ≥ 20 mg/day. At the time of enrollment, the patient's disease must be in remission and their prednisone dose ≥ 5 mg/day and ≤ 20 mg/day. If the patient is taking an immunosuppressive medication agent other than prednisone (maintenance agent) then the maintenance agent must be at a stable dose for 1 month prior to enrollment with no plans by the treating physician to change the dose (other than for safety purposes or toxicity) for the duration of the study (through the month 6 visit or early termination). Participants must be age 18 years or above and their treating physician needs to agree that either treatment assignment is standard of care.

All enrolled participants are tapered from their baseline prednisone dose to 5 mg/day and are then randomized using a 1:1 ratio to remain on 5 mg/day or taper their prednisone dose down to 0 mg/day for the duration of the study (approximately 6 months) or until a study endpoint (Figure 1). All study participants are followed for 6 months or until an increase of prednisone dose (after randomization) occurs, whichever comes first. CoE participants have up to 4 study visits, a screening visit (visit 1), a baseline (visit 2), a month 3 visit (visit 3), and a month 6 or flare visit (visit 4), and up to 2 follow-up phone calls from the study coordinator at randomization and at month 1 (randomization and month 1 phone call may be combined if randomization occurs at month 1).

The primary study endpoint outcome at 6 months included: either (1) continued remission; or (2) a relapse, defined as the physician decision to increase glucocorticoids for disease relapse. The secondary study endpoints include: rates of flare subtypes (severe vs nonsevere), time to event (flare), health-related quality of life, safety (adverse events), and protocol performance (compliance, retention, data completeness, timeliness of data entry, and data accuracy).

Figure 1. The Assessment of Prednisone in Remission (TAPIR) study summary for center of excellence (CoE) and patient-centric (PC) arms.



^Prednisone dose between 11 mg and 20 mg/day: subjects can be screened & enrolled.
Prednisone dose between 6 mg and 10 mg/day: subjects can be screened, enrolled, and have a baseline visit.
Prednisone dose of 5 mg/day: subjects can be screened, enrolled, have baseline visit and be randomized.
*Flare visit is the early termination visit.
+At month 1, Coordinator will call subject to confirm prednisone dose.

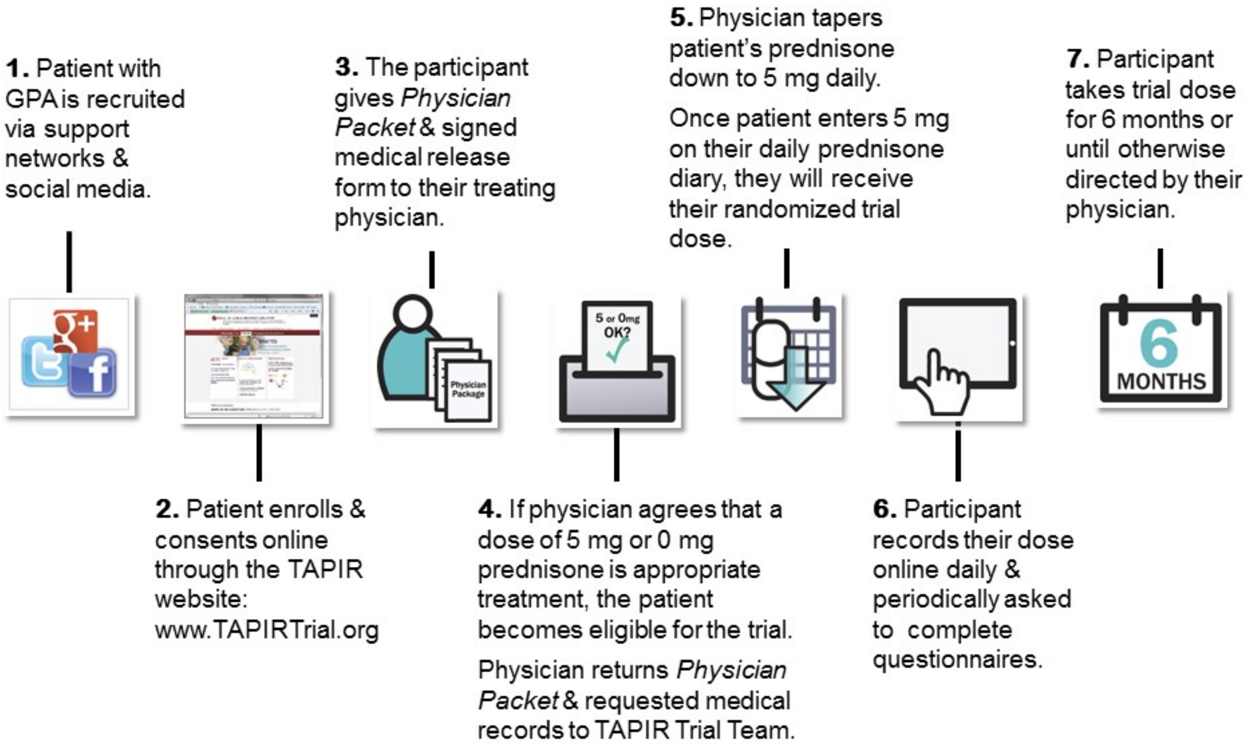
Patient-Centric Recruitment

The VCRC Contact Registry, social media websites, and the Vasculitis Foundation, the largest patient advocacy group for vasculitis, are utilized to direct patients to the study public website (Figure 2). The study's public website provides information including inclusion or exclusion criteria, requirements for participation in the study, the study design of tapering of prednisone to 0 mg/day or 5 mg/day, and whom to contact with questions.

The public TAPIR website also contains an interactive informed consent form to enroll in the study. Potential participants are presented with a video about the study that explains the goals of the research and the risks and benefits of the study. Participants are able to access this video continually during the consenting process and after enrollment. The participant is able to contact study staff through social media or other options including phone or email. Once enrolled in the study, the participant has access to a study website that is personalized for each participant. Participants can keep track of their progress in the study, access the Web-based consent form, and access study forms. Participants receive their treatment randomization assignment via the participant website as well as a physician packet that the participant takes to his or her treating physician explaining the research study and the participant's involvement in the study. The physician has the option to agree to be contacted by study staff to answer questions about their role in this research study.

In order for the participant to be eligible for the study and to establish that the treating physician is not engaged as a researcher in the study, as defined by the US Office for Human Research Protections, the treating physician needs to agree that either a prednisone dose of 5 mg/day or 0 mg/day is standard of care; such agreement is thus consistent with the treating physician providing routine care. A protocol oversight management team provides independent review of a patient's medical records provided by the treating physician to determine if the participant meets eligibility requirements. Participants taper their prednisone dose under the guidance of their own treating physician. Once the participant reaches a dose of prednisone of 5 mg/day, the participant reengages the website and is randomized to continue prednisone at 5 mg/day or taper prednisone to 0 mg/day. The participant's treating physician is then notified of the participant's randomized dose. Participants are to be followed for 6 months from randomization.

Figure 2. Patient-centric (PC) arm study flow.



1. Patient with GPA is recruited via support networks & social media.

2. Patient enrolls & consents online through the TAPIR website: www.TAPIRTrial.org

3. The participant gives *Physician Packet* & signed medical release form to their treating physician.

4. If physician agrees that a dose of 5 mg or 0 mg prednisone is appropriate treatment, the patient becomes eligible for the trial.

Physician returns *Physician Packet* & requested medical records to TAPIR Trial Team.

5. Physician tapers patient's prednisone down to 5 mg daily.

Once patient enters 5 mg on their daily prednisone diary, they will receive their randomized trial dose.

6. Participant records their dose online daily & periodically asked to complete questionnaires.

7. Participant takes trial dose for 6 months or until otherwise directed by their physician.

Center of Excellence Recruitment Model

Investigators at CoE sites are responsible for implementation of the protocol, including screening of potential participants, enrollment of participants, conduct of the protocol, and adherence to federal and local guidelines for clinical research and the protection of human participants (IRB approval). Site-specific CoE study coordinators are responsible for managing the day-to-day operations and implementation of the TAPIR protocol, including the completion of all relevant documentation and record-keeping; scheduling participants for study visits and maintaining the visit calendar; collecting the prednisone study diary; following principles of Good Clinical Practice; submitting the protocols and consent forms to the institutional review or ethics boards; and assisting with patient education and training.

Recruitment occurs through the clinical practices of each site. Participants are also recruited, as needed, via mailings to appropriate clinicians in the investigators' catchment area. Details of the goals of the research and the risk and benefits of the protocol are reviewed with each potential study participant. The consenting process is documented in the study chart.

An optional part of the study was participation in a 21-question knowledge assessment to compare the Web-based PC-arm based informed consent with the standard administration of an informed consent in the CoE. The knowledge assessment contained questions regarding the purpose of the study, eligibility, treatment arm dosing, risks, and communication with the study team or the individual's treating physician.

Qualitative Interviews

Semistructured qualitative interviews were conducted with patients with GPA in both the CoE and PC arms. Patients were interviewed following the consent process and at any of 3 study endpoints (6-month completion, flare related drop-off, nonflare related drop-off). Regardless of study arm, patient interview domains included: (1) factors affecting the decision to enroll, (2) motivation for participation, (3) perception of the recruitment or consent process, (4) expectations for participation, including risks and benefits, and (5) understanding and comfort with tapering period or process. Additional questions were added for participants in both study arms that did not reach the 6-month completion endpoint.

Participant interviews were conducted over the phone. With the participant's permission, all interviews were audio recorded and professionally transcribed. All deidentified qualitative data was entered into NVivo 10 (QSR International), qualitative data analysis software, to facilitate analysis. Interviews were analyzed using modified grounded theory techniques. After a close reading of the initial transcripts, a codebook was developed comprised a priori and grounded theory codes. This analytical method allowed the team to code for ideas of particular interest to the study team and to code for ideas that emerged from the participants' words. Twenty percent of all transcripts were coded by 2 coders in order to assess interrater reliability. If there was less than 90% agreement or .6 kappa in coding, the coders and investigators discussed the areas under question and discrepancies were resolved by consensus methods.

Results

Enrollment Findings

The PC and CoE recruitment efforts were launched on February 17, 2014, after receiving IRB approval. The CoE effort was initiated at the 2 sites that had received IRB approval at this point in time. There was a 2-month lapse between the activation of the first 2 CoE sites and the eighth and final CoE site. Distribution of recruiting materials by the Vasculitis Foundation also began on February 17, 2014; this included promotion of the study on their website, creation of a separate TAPIR trial Web page, inclusion in their quarterly newsletter, a webinar on February 15, 2015, and mention at Chapter meetings and the 2015 Vasculitis Foundation symposium. Distribution to the 1458 patients with GPA in the United States enrolled in the Vasculitis Patient Contact Registry began on May 13, 2014, and was repeated in one-to-two month intervals.

Enrollment, as of May 31, 2016, was 49 in the CoE and 10 in the PC arms. Planned enrollment was 3.3 participants per month for each of the arms, whereas the actual enrollment rate was 0.4 and 1.8 participants per month for the PC and CoE arms, respectively. The social media–directed recruitment effort brought 16,094 individuals to the TAPIR trial website over this time period. Of this large group, only 82 (0.5%, 82/16,094) consented to participate in the study (Figure 3). Information on diagnosis and prednisone dosing was provided during registration by 60 of these 82 (73%; 60/82). Self-reported data suggested 47 of the 60 (78%; 47/60) were potentially eligible for enrollment. The distribution of demographic and GPA diagnosis clinical data for participants enrolled is summarized in Table 1.

Table 1. Study participant demographics and medical history.

Demographics		Patient-centric arm (n=10)	Center of excellence arm (n=49)
Age at enrollment (years)			
	Mean	54.8	55.6
	Median	56.5	59
	Range	37-69	21-80
Sex			
	Male	3	25
	Female	7	24
	Not indicated	0	0
Age at diagnosis (years)			
	Mean	47.7	53.1
	Median	47	56
	Range	31-65	20-80
Newly diagnosed versus recurrent disease			
	Newly diagnosed	1	28
	Recurrent disease	9	21
Years since diagnosis			
	Mean	8.4	3.4
	Median	7	2
	Range	1-24	0-23

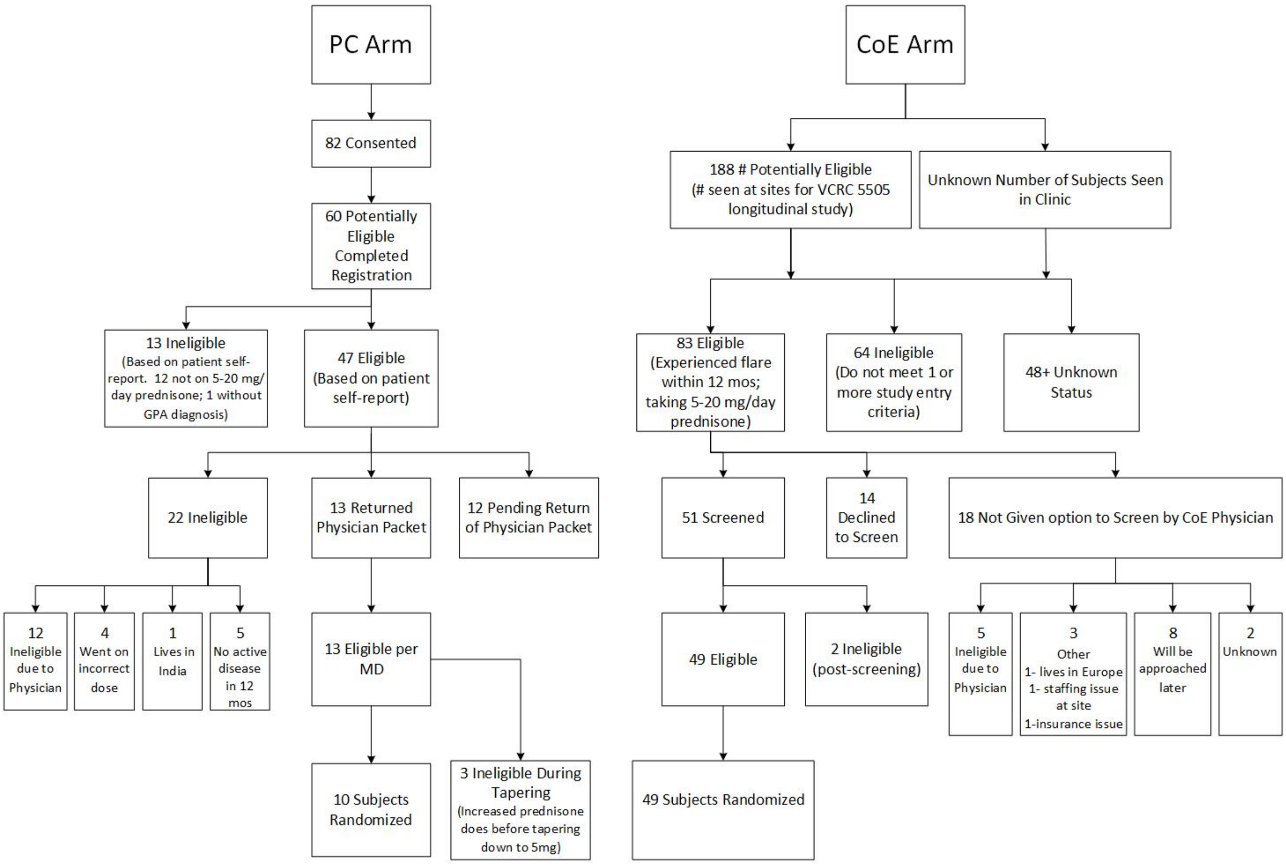
After review of medical records, 64 of the 188 (34%; 64/188) potential CoE participants were found to be ineligible as compared with 13 of 60 (22%; 13/60) in the PC arm (*P*=.10). The 2 approaches toward recruitment were not significantly different in provider acceptance with 18 of the 83 (22%; 18/83) eligible participants on the CoE arm excluded by their treating physician as compared with 12 of 47 (26%; 12/47) on the PC arm (*P*=.78). The actual enrollment of those confirmed eligible by their physician was 10 of 13 (77%; 10/13) and 49 of 51 (96%; 49/51) in the PC and CoE arms (*P*=.11), respectively.

In the PC arm, those who were not eligible were not on the appropriate prednisone dose (n=12) or did not have GPA (n=1). Those who were eligible were requested to contact their treating physician to confirm eligibility and obtain physician concurrence that the treatment plan of either 5 mg/day or 0 mg/day of prednisone was acceptable (n=47). Despite a multipronged approach or reminders and attempted direct contact with treating physicians by study staff, this step posed a significant obstacle with 26% of consenting eligible participants unable to proceed with the study due to the inability to get physician concurrence. Of the 35 physician packets returned, an additional 22 people (63%, 22/35) were found to be ineligible for various reasons (Figure 3), leaving 13 eligible participants. Of those, 10 were randomized to one of the 2 TAPIR treatment arms. By comparison, the CoE sites identified 147 potentially eligible individuals of whom 83 (57%, 83/147) met the TAPIR trial eligibility criteria. 51 of 83 (61%; 51/83) were found to be eligible; 49 of these 51 patients were randomized.

At present, none of the 10 participants (0%) on the PC arm and 2 of 49 (4%; 2/49) participants on the CoE arm have withdrawn from the study before reaching the study end point (*P*>.99). This compares favorably with the anticipated 20% withdrawal rate included in the planning for the trial. One participant on the PC arm and 9 on the CoE are currently on study. As those participants are in follow-up and the randomized trial is still accruing, no data on clinical outcomes is presented.

With 14 of 60 PC and 37 of 50 CoE participants responding, the mean (standard deviation) informed consent knowledge scores, which asked questions about the study purpose, eligibility, expectations regarding study compliance and data reporting, and the ability to withdraw from the study, were 9.8 (88%) and 10.1 (85%), (*P*=.34). In total, 47% of all participants answered more than 91% of the questions correctly. Additionally, 7% of PC participants and 8% of CoE participants answered fewer than 70% of the questions correctly.

Figure 3. The Assessment of Prednisone in Remission (TAPIR) trial consort diagram.



Qualitative Findings

A total of 19 qualitative interviews were conducted with 14 participants recruited from the CoE arm and 5 from the PC arm. Qualitative data illustrated factors driving study enrollment and retention. Participants described a range of inwardly- and outwardly-directed rationales for participating in the study. Outwardly-directed rationales included the stated desire to help others with vasculitis and to further scientific knowledge about the treatment of vasculitis. More inwardly-focused drivers included a sense of a having access to a larger medical network while participating in studies and, specifically for this study, a means of expediting their process of stopping or reducing their prednisone.

Information sources related to study enrollment varied among participants. Some participants described learning about the study through emails or information available on websites (eg, Facebook). Others described finding out about the study through their treating physician. The described information sources were also key factors in patients' self-assessment of eligibility and appropriateness of testing the study question. Participants described challenges related to study enrollment and some of the specific aspects of the tested study designs. Negotiating risk was a factor in participants' decision to enroll in the study. Though some participants did not identify risk in study enrollment, others voiced concerns over confidentiality, randomization to an undesired study arm, and concerns over associated risk of flare or sickness. Primary factors driving retention included descriptions of study physicians and study coordinators being viewed as trusted sources of information and having availability to address participants' concerns about perceived risks. Multimedia Appendix 1 provides illustrative quotes supporting the summative findings above.

Discussion

Principal Findings

The Web-based PC approach to recruiting participants into this study was clearly not as successful as the traditional CoE approach. Nonetheless, patients are capable of understanding and correctly appraising whether or not they met the eligibility criteria for the trial and equally astute as to whether their treating physician would agree with the study. Furthermore, there was no difference in the understanding of the trial as reflected in the knowledge surveys of individuals on the PC and CoE arms. In many respects the enrollment yield from the population of potentially eligible participants is the same for the PC or CoE recruitment strategies. The major difference that accounts for the larger number of those enrolled in the PC arm stems from the higher number of eligible subjects.

The ability to get community physicians to facilitate the enrollment process by certifying that either of the TAPIR treatments (prednisone 0 or 5 mg/day) was consistent with the standard of care proved to be problematic. Although eliminating this problem would not have made the accrual rates equal between the arms, it would likely have substantially positively contributed to the accrual numbers. This barrier existed despite the fact that implementation of the PC arm did not require treating community physicians to be involved in the IRB process, nor any more effort other than acknowledging that the treatment choices were medically appropriate, and releasing the medical records. The qualitative findings illustrated that some participants described the physician approval as a both a barrier to enrollment for interested participants and a potential deterrent. Even though not considered part of the study team, PC arm patients ascribed much weight to their treating physician's opinion about participation and role in or perceived burden of establishing eligibility.

Web-based social media was successful in mobilizing a substantial number of individuals to the study website. Qualitative data supported the intention that these methods would be an efficient way to reach and motivate large numbers of people. The qualitative data suggested that there was interest among participants in the study question as well as an overarching desire to advance knowledge related to the treatment of vasculitis. Others described enrollment as a means of getting off of prednisone where they may not have if not participating in the study. It is possible that increased success of PC recruitment could have been seen with a more broadly applicable or alternative study question. Alternatively, age has been shown to be a moderating factor in the appeal of direct-to-consumer marketing, with older adults relying more on their providers rather than the marketed message [19] which implies that the direct PC appeal of the TAPIR trial might have also have been muted due to the age distribution of this disease.

The prospect of reduced cost of PC recruitment as compared with CoE recruitment, with its need for IRB approvals, continues to make PC recruitment highly appealing. For this particular study question the PC approach was not successful in yielding an accrual rate that could support achievement of the target sample size in a reasonable amount of time and it was substantially less successful than the recruitment achieved by the CoE. Notably, the CoE accrual has also been less than anticipated at the onset of the trial, supporting the conclusion that this trial was a challenge for recruitment in general. The experience of exploring the PC approach in this study yielded several interesting observations and methods that should be evaluated in other clinical settings to get a better sense of the potential of these new methods for clinical trial conduct.

Conclusions

This study did not suffer from the same issues as the Research on Electronic Monitoring of OAB Treatment Experience (REMOTE) [20] study, which also attempted Web-based recruitment and enrollment in 2011, but it shared a similar fate. The potential benefit of direct-to-patient recruitment remains to be realized and it may be that the characteristics of the best clinical setting and target population are yet to be found.

Acknowledgments

This project was supported by the National Heart, Lung, and Blood Institute (R01HL115041). The Vasculitis Clinical Research Consortium (VCRC) has received support from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (U54 AR057319), the National Center for Research Resources (U54 RR019497), the Office of Rare Diseases Research (ORDR), and the National Center for Advancing Translational Science (NCATS). The VCRC is part of the Rare Diseases Clinical Research Network (RDCRN), an initiative of ORDR and NCATS.The RDCRN Data Coordinating Center (U01TR001263) is also supported by ORDR, NCATS.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Qualitative themes and quotations from participant interviews.

[PDF File (Adobe PDF File), 137KB - jmir_v19i2e50_app1.pdf]

References

1. Cheng SK, Dietrich MS, Dilts DM. A sense of urgency: evaluating the link between clinical trial development time and the accrual performance of cancer therapy evaluation program (NCI-CTEP) sponsored studies. Clin Cancer Res 2010 Nov 15;16(22):5557-5563 [FREE Full text] [doi: 10.1158/1078-0432.CCR-10-0133] [Medline: 21062929]
2. Ramsey S, Scoggins J. Commentary: practicing on the tip of an information iceberg? Evidence of underpublication of registered clinical trials in oncology. Oncologist 2008 Sep;13(9):925-929 [FREE Full text] [doi: 10.1634/theoncologist.2008-0133] [Medline: 18794216]
3. Carlisle BC, Kimmelman J, Ramsay T, MacKinnon N. Unsuccessful trial accrual and human subjects protections: an empirical analysis of recently closed trials. Clin Trials 2015 Feb;12(1):77-83 [FREE Full text] [doi: 10.1177/1740774514558307] [Medline: 25475878]
4. Richesson RL, Lee HS, Cuthbertson D, Lloyd J, Young K, Krischer JP. An automated communication system in a contact registry for persons with rare diseases: scalable tools for identifying and recruiting clinical research participants. Contemp Clin Trials 2009 Jan;30(1):55-62 [FREE Full text] [doi: 10.1016/j.cct.2008.09.002] [Medline: 18804556]

5. Donohue JM, Cevasco M, Rosenthal MB. A decade of direct-to-consumer advertising of prescription drugs. N Engl J Med 2007 Aug 16;357(7):673-681. [doi: 10.1056/NEJMsa070502] [Medline: 17699817]
6. Greene JA, Kesselheim AS. Pharmaceutical marketing and the new social media. N Engl J Med 2010 Nov 25;363(22):2087-2089. [doi: 10.1056/NEJMp1004986] [Medline: 21105789]
7. Ventola CL. Direct-to-consumer pharmaceutical advertising: therapeutic or toxic? P T 2011 Oct;36(10):669-684 [FREE Full text] [Medline: 22346300]
8. Smith SK, Selig W, Harker M, Roberts JN, Hesterlee S, Leventhal D, et al. Patient engagement practices in clinical research among patient groups, industry, and academia in the united states: a survey. PLoS One 2015;10(10):e0140232 [FREE Full text] [doi: 10.1371/journal.pone.0140232] [Medline: 26465328]
9. Paul J, Seib R, Prescott T. The Internet and clinical trials: background, online resources, examples and issues. J Med Internet Res 2005;7(1):e5 [FREE Full text] [doi: 10.2196/jmir.7.1.e5] [Medline: 15829477]
10. Marks R, Bristol H, Conlon M, Pepine CJ. Enhancing clinical trials on the internet: lessons from INVEST. Clin Cardiol 2001 Nov;24(11 Suppl):V17-V23 [FREE Full text] [Medline: 11712772]
11. Köpcke F, Prokosch H. Employing computers for the recruitment into clinical trials: a comprehensive systematic review. J Med Internet Res 2014 Jul 01;16(7):e161 [FREE Full text] [doi: 10.2196/jmir.3446] [Medline: 24985568]
12. Valdez RS, Guterbock TM, Thompson MJ, Reilly JD, Menefee HK, Bennici MS, et al. Beyond traditional advertisements: leveraging Facebook's social structures for research recruitment. J Med Internet Res 2014;16(10):e243 [FREE Full text] [doi: 10.2196/jmir.3786] [Medline: 25348050]
13. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. J Med Internet Res 2007;9(2):e11 [FREE Full text] [doi: 10.2196/jmir.9.2.e11] [Medline: 17513282]
14. Bull SS, Vallejos D, Levine D, Ortiz C. Improving recruitment and retention for an online randomized controlled trial: experience from the Youthnet study. AIDS Care 2008 Sep;20(8):887-893. [doi: 10.1080/09540120701771697] [Medline: 18777217]
15. Ross S, Grant A, Counsell C, Gillespie W, Russell I, Prescott R. Barriers to participation in randomised controlled trials: a systematic review. J Clin Epidemiol 1999 Dec;52(12):1143-1156. [Medline: 10580777]
16. Walsh M, Merkel PA, Mahr A, Jayne D. Effects of duration of glucocorticoid therapy on relapse rate in antineutrophil cytoplasmic antibody-associated vasculitis: a meta-analysis. Arthritis Care Res (Hoboken) 2010 Aug;62(8):1166-1173 [FREE Full text] [doi: 10.1002/acr.20176] [Medline: 20235186]
17. Krischer JP, Gopal-Srivastava R, Groft SC, Eckstein DJ, Rare Diseases Clinical Research Network. The Rare Diseases Clinical Research Network's organization and approach to observational research and health outcomes research. J Gen Intern Med 2014 Aug;29(Suppl 3):S739-S744 [FREE Full text] [doi: 10.1007/s11606-014-2894-x] [Medline: 25029976]
18. Hunder GG, Arend WP, Bloch DA, Calabrese LH, Fauci AS, Fries JF, et al. The American College of Rheumatology 1990 criteria for the classification of vasculitis: introduction. Arthritis Rheum 1990 Aug;33(8):1065-1067. [Medline: 2390119]
19. Ball JG, Manika D, Stout PA. The moderating role of age in responses to direct-to-consumer prescription drug advertising. J Health Commun 2016;21(1):12-32. [doi: 10.1080/10810730.2015.1023960] [Medline: 26312583]
20. Covington D, Veley K. Appliedclinicaltrialsonline. 2015 Feb 01. The remote patient-centered approach in clinical research URL: http://www.appliedclinicaltrialsonline.com/print/287791?page=full [accessed 2017-01-12] [WebCite Cache ID 6nRouDXrl]

Abbreviations

CoE: center of excellence
GPA: granulomatosis with polyangiitis
PC: patient-centric
RCT: randomized controlled trial
RDCRN: Rare Diseases Clinical Research Network
TAPIR: The Assessment of Prednisone in Remission (TAPIR)
VCRC: Vasculitis Clinical Research Consortium (VCRC)

Edited by G Eysenbach; submitted 12.10.16; peer-reviewed by M Harker, L Deblieck, C Polanco; comments to author 30.11.16; revised version received 18.01.17; accepted 21.01.17; published 28.02.17.

Please cite as:
Krischer J, Cronholm PF, Burroughs C, McAlear CA, Borchin R, Easley E, Davis T, Kullman J, Carette S, Khalidi N, Koening C, Langford CA, Monach P, Moreland L, Pagnoux C, Specks U, Sreih AG, Ytterberg S, Merkel PA, Vasculitis Clinical Research Consortium
Experience With Direct-to-Patient Recruitment for Enrollment Into a Clinical Trial in a Rare Disease: A Web-Based Study
J Med Internet Res 2017;19(2):e50
URL: http://www.jmir.org/2017/2/e50/
doi:10.2196/jmir.6798
PMID:28246067

Original Paper

Association Between Internet Use and Body Dissatisfaction Among Young Females: Cross-Sectional Analysis of the Canadian Community Health Survey

Allison Carter[1,2], MPH; Jamie I Forrest[3], MPH; Angela Kaida[1], PhD

[1]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada
[2]Epidemiology and Population Health Program, British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada
[3]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Allison Carter, MPH
Epidemiology and Population Health Program
British Columbia Centre for Excellence in HIV/AIDS
St. Paul's Hospital
608-1081 Burrard Street
Vancouver, BC, V6Z 1Y6
Canada
Phone: 1 604 558 6688
Fax: 1 604 559 9800
Email: allison_carter@sfu.ca

Abstract

Background: Recent research suggests Internet exposure, including Facebook use, is positively correlated with body dissatisfaction, especially among girls and young women. Canada has one of the highest Internet access rates in the world, yet no previous study has examined this relationship using nationally representative data.

Objective: Our objective was to evaluate the relationship between Internet use and body dissatisfaction among a national, population-based sample of Canadian females 12-29 years of age.

Methods: We used cross-sectional data from the Canadian Community Health Survey 2011-2012. Body dissatisfaction was measured using a 5-point Likert scale and defined as “very dissatisfied/dissatisfied” with one’s body. The explanatory variable was time spent using the Internet per week in the past 3 months, ranging from none/<1 hour to >20 hours. We used multinomial logistic regression to investigate whether greater Internet use was associated with increasing odds of being very dissatisfied/dissatisfied, neutral, or satisfied with one’s body, using very satisfied as the referent. Probability survey sampling weights were applied to all analyses.

Results: Of 2983 included participants, sampled to represent 940,786 young Canadian females, most were 20-29 years old (61.98%) and living in households with an annual income Can $80,000 or more (44.61%). The prevalence of body dissatisfaction was 14.70%, and 25- to 29-year-olds were more likely than 12- to 14-year-olds to be very dissatisfied or dissatisfied with their body (20.76% vs 6.34%). Few (5.01%) reported none/<1 hour of Internet use, over half (56.93%) reported 1-10 hours, and one-fifth (19.52%) reported spending >20 hours online per week. Adjusting for age and income, the odds of being very dissatisfied/dissatisfied, relative to very satisfied, were greater in the highest versus lowest Internet use group (adjusted odds ratio [AOR] 3.03, 95% CI 1.19-7.70). The AORs for this level of body dissatisfaction increased across increasing levels of Internet use, ranging from 0.88 (95% CI 0.35-2.21) to 3.03 (95% CI 1.19-7.70). Additionally, those who spent 11-14 hours online were more likely to be neutral (AOR 3.66, 95% CI 1.17-11.45) and those who spent 15-20 hours online were more likely to be neutral (AOR 4.36, 95% CI 1.18-16.13) or satisfied (AOR 2.82, 95% CI 1.14-7.01) with their bodies, relative to very satisfied, compared with those spending no time or <1 hour online.

Conclusions: A substantial proportion of Canadian females 12-29 years of age spent large amounts of time (>20 hours) on the Internet each week, and body dissatisfaction was significantly more likely among this group. Those who spent 11-20 hours online were also more likely to be less satisfied with their bodies. Efforts are needed to support girls and young women to achieve and maintain a positive body image in today’s digital age.

(J Med Internet Res 2017;19(2):e39) doi:10.2196/jmir.5636

KEYWORDS

body satisfaction; Internet use; girls; young women; Canada

Introduction

Body dissatisfaction, defined by a subjective negative view of one's body weight and shape [1], is an important public health issue globally and in Canada, especially among girls and young women. An international survey of 11- to 15-year-old adolescents in 24 regions in Europe, the United States, and Canada found a high prevalence of body dissatisfaction across all countries, with variation seen by sex, age, and weight [2]. Among Canadians included in the study, 43% of girls (and 27% of boys) reported feeling dissatisfied with their body, and being older and overweight was positively correlated with body dissatisfaction [2]. In another longitudinal study of Canadian adolescents in Quebec, 57.1% of adolescent girls desired a thinner body at age 14 years, with that prevalence increasing to 65.8% at age 18 years; this is compared with 44.0% among adolescent boys at baseline, which remained unchanged 4 years later (39.5%; *P*=.17) [3]. Similarly high rates of body dissatisfaction are also seen among women during young adulthood, a time when social comparisons intensify [4,5]; research demonstrates that such comparisons, not just with peers but also with sociocultural images of thinness and attractiveness (eg, models and actors), can be detrimental to body image [6]. Understanding and mitigating risk factors for body dissatisfaction among females at an early age is critical given the high number of associated poor physical and mental health outcomes over the longer term, including low self-esteem [7], depression and suicidal ideation [8], disordered eating and exercise patterns [9], and "risky" sexual behaviors [10].

A large body of evidence suggests that sociocultural influences that equate female beauty with thinness play a powerful role in the development of body dissatisfaction among girls and young women [5,11,12]. Two reviews of experimental and correlational studies, including one meta-analysis, demonstrated that exposure to media such as television and magazines, which have ubiquitous images, messages, and advertisements promoting the ideal that women need to be thin, is associated with higher odds of body dissatisfaction [13,14]. More recently, interest has turned to the role of the Internet. Although young people use the Internet for multiple purposes, there is evidence that many of the websites and advertisements aimed at and accessed by girls and young women focus heavily on aspects of idealized female beauty (eg, health, fitness, fashion, makeup, and celebrities) [11]. Recent campaigns such as The Dove Campaign for Real Beauty [15] and Vancouver's Raw Beauty Talks [16], which feature raw portraits of diverse women of varying sizes without makeup and filters, are attempting to alter this narrow picture of what beauty looks like. Nevertheless, the digital mainstream—coupled with the recent and rapid rise of social media platforms, such as Facebook and Instagram, that allow for greater appearance comparisons with people's best photos of themselves (or "highlight reels") [17]—presents tremendous challenges for women wanting to achieve or maintain a healthy body image. Indeed, recent research suggests that Internet exposure, including Facebook use [18-21], is positively correlated with body dissatisfaction.

Canada has one of the highest Internet access rates in the world, with over 85% of households having access to the Internet at home, and the average Internet user spending over 36 hours online every month, outside of work and school [22]. This is in contrast to roughly 40% of households in 2000 [23]. Additionally, it is estimated that 20 million Canadians now have a social networking account, with 18.5 million using Facebook [24], launched in 2004. For girls and young women in Canada, growing up in today's digital age may exacerbate their risk for body dissatisfaction. However, to our knowledge, no previous study has evaluated the relationship between Internet use and body dissatisfaction in this population in Canada, nor used a nationally representative sample. Further, among the few international studies doing pioneering work in this area [18-21], none have examined a possible dose-response relationship. The primary purpose of this study was to answer the following research question: What is the relationship between amount of Internet use and body dissatisfaction among a national, population-based sample of Canadian female adolescents and young adults aged 12-29 years? The main hypothesis was that greater amounts of Internet use would be associated with increasing odds of body dissatisfaction.

Methods

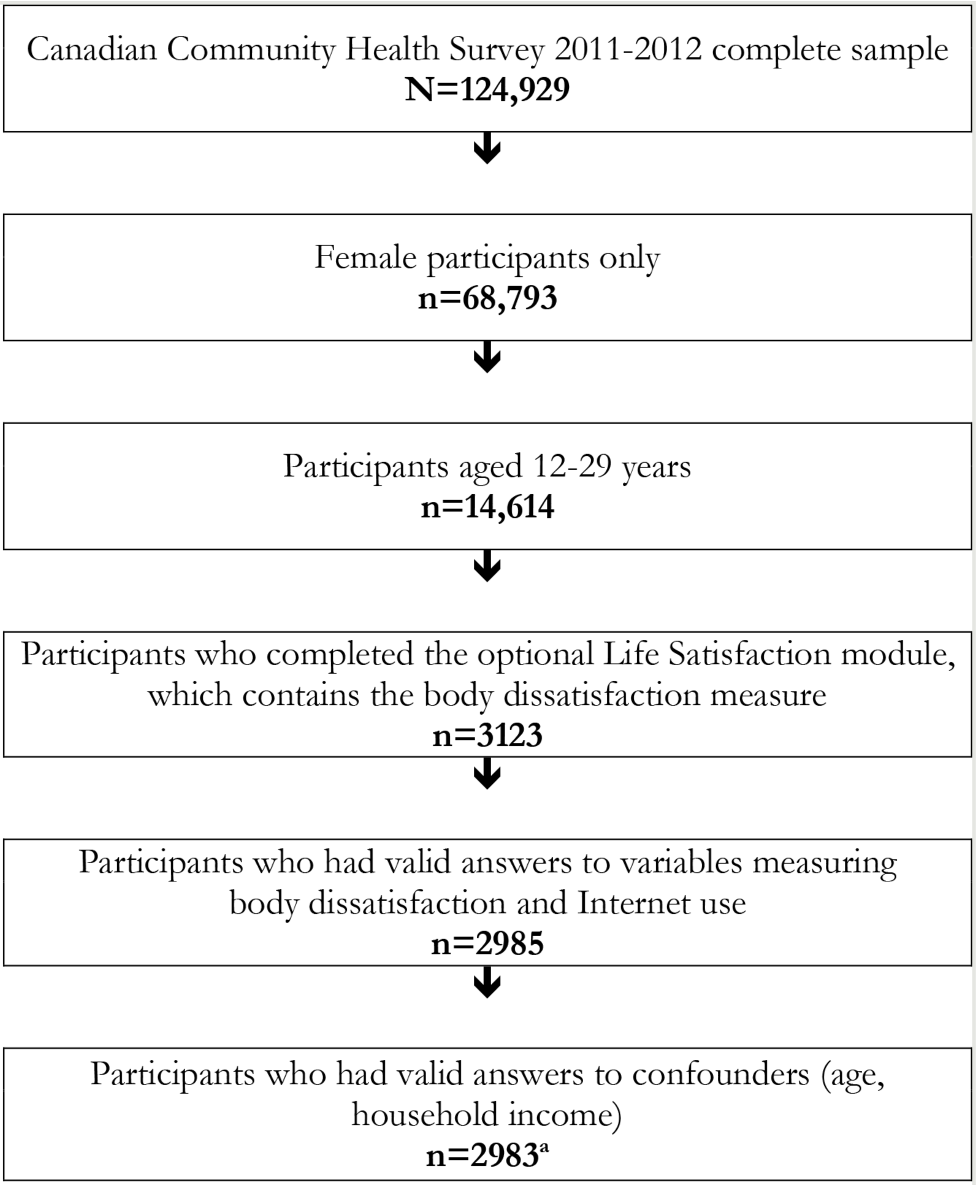
Study Design

This study used data from the Public Use Microdata Files of the Canadian Community Health Survey (CCHS) 2011-2012 [25]. The CCHS is an annual population-based, cross-sectional survey conducted by Statistics Canada that collects information about general health status, personal health behaviors, use of health care services, and various social determinants of health for the purposes of health surveillance and population health research to improve the health of Canadians. The target population is individuals aged 12 years or older from all provinces, territories, and health regions residing in private households. Individuals working in the Canadian Armed Forces or those living in institutions, on First Nations' reserves, or in certain remote areas of northern Ontario and Québec are excluded. Each annual CCHS cycle uses a complex, multistage sampling strategy to randomly select a sample of households over six 2-month collection periods, from January to December. Among sampled households, 1 individual is invited to participate in the survey, based on selection probabilities that consider age and household composition. The sampling strategy is also designed to oversample young people (12-19 years). Participants complete interviewer-administered, computer-assisted questionnaires, either in person or by telephone, that last approximately 60 minutes. A detailed description of the CCHS methodology is available from Statistics Canada [25].

Study Population and Final Analytic Sample

In the CCHS 2011-2012, a total of 124,929 individuals participated out of 144,000 sampled households, representing a response rate of 86.76%. The overall sample was representative of 97% of the Canadian population aged 12 years or older. The participants included in the analysis for this study were female adolescents and young adults aged 12 to 29 years who reported valid responses to the primary outcome measure, explanatory variable, and confounders, described in the section below. As Figure 1 shows, 14,614 females aged 12 to 29 years participated in the CCHS 2011-2012. Of this total, 3123 were asked to complete the optional Satisfaction with Life module, which included perceptions of body dissatisfaction. Of these 3123 eligible adolescent girls and young women, we excluded an additional 230 participants due to invalid responses to the main study variables, such as “don’t know,” “prefer not to answer,” or “not asked/not stated” in the case of interviews completed by a proxy respondent (usually the parent for younger respondents). Thus, the final analytic sample included 2983 adolescent girls and young women. Using the survey weights provided by Statistics Canada, we assigned each person a weight to represent her contribution to the total population, taking into account the complex sampling design. With these survey weights applied, this sample represented 940,786 Canadians of this age and sex.

Figure 1. Flow chart of sample selection.[a]Final sample represents 940,786 Canadian adolescent girls and young women (weighted frequency).



Study Variables

The primary outcome was body dissatisfaction, measured using the following question: “How satisfied are you with the way your body looks?” Possible responses were “very satisfied,” “satisfied,” “neither satisfied nor dissatisfied,” “dissatisfied,” and “very dissatisfied.” We combined the responses “dissatisfied” and “very dissatisfied” due to low sample size and used them as the event of interest. The main explanatory variable was Internet use, assessed by the following proxy question: “In a typical week in the past 3 months, how much time did you usually spend on a computer, including playing computer games and using the Internet or World Wide Web?” Responses were “none or less than 1 hour,” “1-2 hours,” “3-5 hours," "6-10 hours," "11-14 hours," "15-20 hours," and "more than 20 hours." This time frame does not include work or school, allowing for the estimation of leisure computer use, which is time spent largely online [22]. Nevertheless, this is likely a conservative estimate of Internet use, especially as it excludes mobile use (a major Internet source for this demographic) and includes computer games, which is discussed at length in the limitations section and highlights the challenge of using population-based data administered by national statistics agencies. Also, while this specific CCHS measure has not been previously validated, studies suggest that self-report measures are reliable tools to assess Internet use [26]. Confounders known to be determinants of body dissatisfaction and associated with Internet use [13,27] considered in the analyses were age (12-14, 15-17, 18-19, 20-24, and 25-29 years) and annual household income (in Can $: <20,000, 20,000-39,999, 40,000-59,999, 60,000-79,999, ≥80,000).

Other confounders we considered were ethnicity and body mass index (BMI). The CCHS categorizes ethnicity as white versus nonwhite according to a predetermined list without consideration of other meaningful dimensions of racial identity that may differentially affect body positivity. As such, we did not include ethnicity as a confounder in this analysis, and we feel that controlling for household income blocked some of the effects of ethnicity. Indeed, bivariable analyses revealed a strong association between these 2 variables, with increasing proportions of participants identifying as white, moving from the lowest to the higher income categories: 50.76%, 56.80%, 56.07%, 65.53%, and 74.17%, respectively. BMI was a strong predictor of body satisfaction, so much so that over 90.20% of those who were *not* overweight/obese were very satisfied with their bodies (n=468), with a small cell size in this outcome level for those who were overweight/obese (n=48). Inclusion of BMI in the adjusted multivariable model significantly increased the magnitude of the effect estimates of Internet use on body dissatisfaction and widened the variance around those estimates considerably. On one hand, this is an indicator of a strong confounder, but on another hand, it could be overadjustment, since BMI is highly correlated with body satisfaction. Importantly, the addition of BMI did not change the overall conclusions of the study. Further, in adding BMI, we lost 310 participants from the sample due to missing or invalid responses to this measure. Given these issues (eg, collinearity with body satisfaction, inflation of estimates and variance, and reduced sample size), as well as the limitations associated with the broad proxy measure of Internet use provided by the CCHS, we decided to take a conservative approach and exclude BMI from the final model, since it raised more issues regarding the statistical and analytical models than it solved conceptually.

Analysis Plan

We calculated descriptive statistics using frequencies (n) and weighted percentages (%) to provide baseline characteristics, and the prevalence and patterns of body dissatisfaction and Internet use for the overall study sample. Bivariable analyses were conducted of the explanatory variable and confounders by the outcome measure. Associations were tested using Pearson chi-square test statistic. Binomial logistic regression [28] analyses were used to produce unadjusted estimates (crude odds ratios [ORs] and 95% CIs) of the association between body dissatisfaction and Internet use, as well as for the association between the confounders and the outcome. On review of the bivariable results, we used a multinomial logistic regression model [29], adjusting for confounders, to investigate whether greater amounts of Internet use were associated with increased odds of being very dissatisfied/dissatisfied, neutral, or satisfied with one's body, using very satisfied as the referent, and report adjusted ORs (AORs) and 95% CIs.

For all models, age shifted the estimates by a significant amount (ie, >40%), while we observed only minor adjustments when adding household income to the model. Model fit statistics were better with income added (ie, lower Akaike information criterion) [30]; given this, and literature that suggests income may be an important confounder [27], we kept income in the model. Sensitivity analyses were also conducted to examine whether the number of variables in the model or an uneven distribution of cell sizes may have been contributing to the variability around the estimates; accordingly, we treated household income as a continuous measure and collapsed the 2 lowest and 2 highest categories of Internet use. The AORs and 95% CIs remained similar, and we observed a slightly poorer model fit (ie, higher Akaike information criterion); thus, we retained the original model.

All analyses were conducted using SAS version 9.3 for Mac (SAS Institute Inc). As mentioned, we weighted respondent data to account for the nonrandom sampling scheme and uneven probability of selection into the survey sample, using probability survey sampling weights developed and provided by Statistics Canada. These weights were applied to all analyses in order to produce more precise point estimates and variances (for details of the weighting procedures: [31]). Ethics approval for this study was covered under the publicly accessible data clause (Item 1.3.1) of the University of British Columbia's Policy #89 on Research and Other Studies Involving Human Subjects.

Results

Baseline Characteristics

Of the 2983 adolescent girls and young women included in this study, most were 20 to 29 years of age (61.98%), versus less than 20 years, and lived in households where the annual income was Can $80,000 or greater (44.61%) (the highest level available for self-report in CCHS). Table 1 shows the baseline characteristics of this population as well as the patterns of Internet use and confounders by different levels of body dissatisfaction or satisfaction.

Table 1. Baseline characteristics and bivariable associations with body satisfaction and dissatisfaction among Canadian adolescent girls and young women aged 12-29 years: Canadian Community Health Survey 2011-2012 (n=2983).

Variables	Total	Body satisfaction				
	n (%)[a]	Very dissatisfied/ dissatisfied (n=440, 14.70%) n (%)[a]	Neutral (n=349, 12.82%) n (%)[a]	Satisfied (n=1604, 53.44%) n (%)[a]	Very satisfied (n=590, 19.04%) n (%)[a]	*P* value
Internet use (per week in the past 3 months), in hours						<.001
None/<1	159 (5.01)	33 (6.1)	14 (3.7)	79 (4.4)	33 (6.8)	
1-2	529 (16.32)	64 (13.4)	63 (13.5)	270 (16.8)	132 (19.0)	
3-5	556 (17.48)	61 (11.7)	64 (17.2)	324 (19.5)	107 (16.4)	
6-10	749 (23.13)	98 (18.0)	80 (17.4)	415 (24.4)	156 (27.3)	
11-14	358 (11.97)	53 (11.1)	46 (18.1)	194 (11.5)	65 (9.7)	
15-20	167 (6.55)	33 (6.8)	21 (8.2)	86 (6.8)	27 (4.6)	
>20	465 (19.52)	98 (33.0)	61 (21.9)	236 (16.4)	70 (16.3)	
Age, in years						<.001
12-14	473 (13.24)	24 (5.7)	32 (9.4)	240 (11.8)	177 (25.6)	
15-17	517 (15.38)	66 (13.6)	41 (10.5)	284 (15.6)	126 (19.3)	
18-19	325 (9.40)	33 (7.6)	34 (8.5)	204 (10.2)	54 (9.1)	
20-24	726 (30.68)	120 (28.9)	91 (37.0)	394 (30.7)	121 (27.7)	
25-29	942 (31.30)	197 (44.2)	151 (34.5)	482 (31.6)	112 (18.3)	
Annual household income, in Can $						.09
<20,000	210 (7.89)	33 (7.4)	26 (6.7)	113 (7.5)	38 (10.3)	
20,000-39,999	505 (16.72)	83 (15.8)	52 (18.2)	272 (17.3)	98 (14.8)	
40,000-59,999	487 (15.63)	80 (16.4)	63 (18.9)	268 (16.6)	76 (10.2)	
60,000-79,999	471 (15.14)	68 (18.8)	68 (18.7)	255 (14.4)	80 (11.8)	
≥80,000	1310 (44.61)	176 (41.7)	140 (37.3)	696 (44.2)	298 (52.9)	

[a] Column percentages are shown. Percentages have survey weights applied.

Patterns of Internet Use and Confounders by Body Dissatisfaction

The overall prevalence of body dissatisfaction was 14.70%, and 25- to 29-year-olds were more likely than 12- to 14-year-olds to be very dissatisfied or dissatisfied with their body (20.76% versus 6.34%; data not shown). The majority of the sample reported satisfaction with their body at the satisfied (53.44%) or very satisfied level (19.04%). With respect to Internet use, we observed considerable variation in the time participants spent online in a typical week in the past 3 months: few (5.01%) reported none/less than 1 hour of Internet use, over half (56.93%) reported 1-10 hours, 18.52% reported 11-20 hours, and one-fifth (19.52%) reported more than 20 hours online per week. Examining Internet use by the outcome revealed that 33.0% of those who were very dissatisfied or dissatisfied with their body spent more than 20 hours online per week compared with only 16.3% of those who were very satisfied with their body (*P*<.001). In terms of age, among dissatisfied individuals, 44.2% were in the oldest age strata and 5.7% were in the youngest, while the pattern was reversed for very satisfied individuals (*P*<.001).

Association Between Body Dissatisfaction and Internet Use

Table 2 presents the unadjusted ORs and AORs with 95% CIs of reporting increasing levels of body dissatisfaction by increasing amounts of Internet use. The unadjusted odds of being very dissatisfied or dissatisfied with one’s body, relative to very satisfied, were more than 2 times greater in the highest Internet use group (>20 hours) than in the lowest Internet use group (none/<1 hour) (OR 2.27, 95% CI 0.92-5.56). After adjusting for age and household income, the magnitude of the association between Internet use and body dissatisfaction was stronger (AOR 3.03, 95% CI 1.19-7.70), and the 95% CI excluded the null value of ‘1’ and reflected a range of effects, from moderate to more than a 7-fold increase in the odds of body dissatisfaction. It is noteworthy that the AORs for this outcome level of body dissatisfaction increased across increasing levels of Internet use, ranging from 0.88 (95% CI 0.35-2.21) to 3.03 (95% CI 1.19-7.70). The 95% CIs overlapped and included the null until the maximum level of Internet use; however, the gradual increase in point estimates is noteworthy and may suggest a possible dose-response relationship, in accordance with the principles of effect size estimation emphasizing CIs, as opposed to statistical significance testing [32]. For the 2 other outcome levels (neutral and satisfied), the strength of and variability around the adjusted effect estimates was similar, although we observed no pattern of increasing AORs. For instance, the adjusted odds of reporting being neutral, relative to very satisfied, were almost 4 times greater among those accessing the Internet for 11-14 hours than among those reporting no use or less than 1 hour of use (AOR 3.66, 95% CI 1.17-11.45), with even greater effects seen for those online for 15-20 hours (AOR 4.36, 95% CI 1.18-16.13). Lastly, with 2983 participants and using an alpha of .05, there was power greater than 90% to detect the final adjusted result for the association between body dissatisfaction and Internet use (≥20 hours versus none/<1 hour).

Table 2. Multinomial logistic regression results of being very dissatisfied/dissatisfied, neutral, or satisfied with one’s body in reference to very satisfied, as explained by increasing amounts of Internet use among Canadian adolescent girls and young women aged 12-29 years: Canadian Community Health Survey 2011-2012 (n=2983).

Internet use (per week in the past 3 months), in hours	Very dissatisfied/dissatisfied		Neutral		Satisfied	
	Unadjusted OR[a] (95% CI)	Adjusted OR[b] (95% CI)	Unadjusted OR (95% CI)	Adjusted OR[b] (95% CI)	Unadjusted OR (95% CI)	Adjusted OR[b] (95% CI)
None/<1	Reference	Reference	Reference	Reference	Reference	Reference
1-2	0.79 (0.33-1.91)	0.88 (0.35-2.21)	1.31 (0.46-3.73)	1.41 (0.48-4.19)	1.36 (0.67-2.76)	1.48 (0.71-3.11)
3-5	0.79 (0.32-1.98)	0.90 (0.34-2.36)	1.93 (0.66-5.66)	2.17 (0.72-6.60)	1.82 (0.87-3.79)	2.05 (0.94-4.47)
6-10	0.74 (0.31-1.72)	0.89 (0.36-2.21)	1.17 (0.42-3.31)	1.36 (0.46-3.99)	1.37 (0.68-2.75)	1.55 (0.74-3.26)
11-14	1.28 (0.52-3.18)	1.46 (0.55-3.84)	3.46 (1.13-10.59)	3.66 (1.17-11.45)[c]	1.83 (0.87-3.86)	1.95 (0.88-4.32)
15-20	1.66 (0.58-4.77)	2.51 (0.84-7.45)	3.34 (0.93-12.02)	4.36 (1.18-16.13)[c]	2.28 (0.93-5.60)	2.82 (1.14-7.01)[c]
>20 hours	2.27 (0.92-5.56)	3.03 (1.19-7.70)[c]	2.49 (0.82-7.47)	2.85 (0.92-8.81)	2.49 (0.82-7.47)	1.76 (0.80-3.86)

[a] OR: odds ratio.

[b] Adjusted for age and household income.

[c] AORs excluding the null.

Discussion

Principal Findings

In this national, population-based sample of Canadian females aged 12 to 29 years, 14.70% reported dissatisfaction with their bodies. A substantial proportion of participants spent large amounts of time (≥20 hours) on the Internet each week (19.52%), and body dissatisfaction was significantly more likely among this group. Further, those who spent 11-20 hours online were also more likely to be less satisfied with their bodies. While the data do not provide clear evidence of a dose-response relationship, these results may suggest a possible threshold effect. This is the first Canadian study to explore these patterns using nationally representative data, and these findings add to the international scientific research in this area.

The prevalence of body dissatisfaction in this study varied from estimates in other Canadian cohorts. A 2001/2002 study of 11-, 13-, and 15-year-old Canadian girls and adolescent girls found a much higher prevalence of body dissatisfaction (43.1%) [2], as did a 4-year longitudinal study of young women in Quebec (57.1% at baseline; 65.8% at follow-up) [3]. In another much larger study in Nova Scotia (n=2159), only 7.3% of girls in the sample reported body dissatisfaction [33]. Variations in the measures used may explain some of these differences. For instance, the latter study used “I like the way I look” as a proxy, coding girls responding with “never/almost never” as having body dissatisfaction and comparing them with all remaining girls reporting “sometimes/often/almost always.” The differences may also reflect differences in sample heterogeneity. Notably, the CCHS included a higher income sample, with nearly 60% of participants living in households with an annual income of at least Can $60,000; this is compared with 40% of participants in the Nova Scotia study. Additionally, Statistics Canada uses a complex sampling design to interview a random, representative sample of households across Canada. Thus, while it is unlikely that participants would be self-selected in the CCHS, this may, however, be a factor in other research using nonrepresentative samples.

The finding that high amounts of Internet use are associated with body dissatisfaction reflects previous research linking media exposure to women’s body image concerns. Two reviews reported moderate effect sizes for the relationship between exposure to multiple forms of media promoting the thin ideal and women's body dissatisfaction [13,14]. To our knowledge, however, only a few other studies have specifically examined the impact of Internet use on body dissatisfaction [18-21]. This relationship was first investigated in a 2010 study of 156 Australian females aged 13-18 years [21], in which those authors found effects of Internet use on several body-related constructs, including internalization of thin ideals, appearance comparison, weight dissatisfaction, and drive for thinness. Interestingly, among the range of Internet sites examined in the Australian study, the most significant correlate of poor body image was time spent on Facebook. This study was repeated among a sample of preteenaged girls aged 10-12 years (n=189) [19], as well as a larger and more diverse cohort of 1087 adolescent girls aged 13-15 in Australia [18]. A similar finding was reported: Facebook use was more strongly correlated with poor body image compared with general Internet use. In 2015, Facebook was again linked to body dissatisfaction in a national sample of New Zealand women and men (n=11,017), across ages [20]. Interestingly, that study's authors reported a cohort effect, with younger women who had grown up with more social media exposure reporting lower body satisfaction overall.

Our findings add to this literature, linking the amount of time spent using the Internet to body dissatisfaction and demonstrating an increasing likelihood of body dissatisfaction by increasing amounts of Internet use. The Internet has become ubiquitous in the lives of most Canadian young people, with continual accessibility being common, especially since the emergence of smartphones [22]. The specific content and platforms accessed online are likely wide in variety and remain unexplored in this study due to the lack of collection of such data in the CCHS. However, research suggests that girls and young women are highly exposed to online images and messages about female beauty [18-21], where the unattainable ideal is often someone who is "tall, young, usually white, has long, flowing hair, is surgically enhanced, blemish-free, and very thin" [34]. Social networking sites add to this milieu, creating a space for girls and young women to readily compare themselves with hundreds of peers on a daily basis. As body image experts highlight, "these comparisons are often made with at least somewhat idealized images (girls are likely to post only photographs showing themselves looking good or doing something 'cool,' and even these can be digitally altered)" [19]. Thus, in our study, we hypothesized that time spent online may increase exposure to these sociocultural forces that promote appearance as key to female worth. As previous research has demonstrated [19], this exposure may be related to the internalization of female beauty ideals, and, in turn, body dissatisfaction.

There is a need for future research to understand how the effect of Internet use may vary by age, sex, and time, among other meaningful characteristics. We conducted post hoc analyses to stimulate such efforts (tables not shown). In adjusted analyses with the entire sampled population (n=25,568), we observed a clearer dose-response relationship (with 4 of 6 AORs excluding the null and tighter 95% CIs: 1.29 (0.97-1.71), 1.28 (0.95-1.71), 1.41 (1.07-1.86), 1.78 (1.26-2.52), 2.46 (1.62-3.74), and 3.31 (2.34-4.69). Among those reporting body dissatisfaction, most were women (66.17%) and, of these, there was a roughly equal distribution of body dissatisfaction (18.84% to 22.26%) across these age strata: <30, 30-39, 40-49, 50-59, and ≥60 years. Thus, while individuals under 30 years of age used the Internet the most (accounting for 53% of those online ≥20 hours per week), sex-stratified inquiries into body dissatisfaction across all ages, not just young people, are needed. Next, we repeated analyses among adolescent boys and young men aged 12-29 years (n=2756). Relative to their female counterparts, body dissatisfaction was lower (6.08% vs 14.70%) and Internet use was higher (32.90% reported using it for ≥15 hours per week vs 26.07%). The adjusted odds of reporting being neutral (vs very satisfied) were 3.53 times greater among males accessing the Internet for 15-20 hours (vs none/<1 hour; 95% CI 1.17-10.69); the comparable best estimate for females was higher (4.36, 95% CI 1.18-16.13). All other effect estimates were not significant. These data may suggest sex differences in the link between Internet use and body dissatisfaction. Lastly, we examined data from 2003 (Cycle 2.1), the first year that questions about body satisfaction and Internet use were added to the CCHS and before the rise of smartphones and Facebook. Among females aged 12-29 (n=1565), the prevalence of body dissatisfaction was the same (14.77%) but Internet use was much lower (6.81% were online ≥15 hours per week vs 26.07% in 2011-2012). Adjusting for age and income, only 2 estimates were >2 and excluded the null, both comparing neutral in reference to very satisfied, including AOR 2.61 (95% CI 1.26-5.44) for those online 3-5 hours and AOR 4.61 (95% CI 1.23-17.26) for those online more than 20 hours. This should be explored further in longitudinal research with the same sample followed in time.

In addition to the analyses mentioned above, we conducted one further post hoc sensitivity analysis stratifying by BMI status to see whether the findings persisted across strata (ie, not overweight/obese vs overweight/obese). The sample sizes were extremely small, given 6 categories of Internet use by 4 categories of body satisfaction by 2 categories of weight, resulting in unstable models, estimates, and 95% CIs. However, the models did converge and provide some evidence of an interaction effect that should be investigated in future studies with sufficient sample size. Briefly, the findings for those not overweight/obese were consistent with those for the original model among the full sample—namely, the strongest effects were seen at the highest levels of Internet use (levels 5 through 7), and the magnitude of the effect varied by level of body satisfaction. The highest estimates were seen for neutral relative to very satisfied, although effect estimates were stronger in the normal BMI stratum, but the 95% CIs were very wide (eg, for the outcome of neutral, comparing 15-20 hours of Internet use vs none, the AOR was 4.36, 95% CI 1.18-16.13 in the original model and 12.98, 95% CI 2.22-76.00 in the not overweight/obese stratum). Thus, our conclusions are largely similar for this group. In contrast, 2 distinctive patterns were noteworthy among the stratum of overweight/obese individuals, although again we caution that the 95% CIs were extremely wide. First, the estimates were at their lowest for the neutral versus very satisfied comparisons across all levels of Internet use (eg, for 15-20 hours vs none, the AOR was 3.04, 95% CI

0.21-44.51, which is much smaller than the figures cited above), and the estimates were at their highest for the very dissatisfied/dissatisfied versus very satisfied comparisons across all levels of Internet use (eg, for ≥20 hours vs none, the AOR was 3.03, 95% CI 1.19-7.70 in the original model, 1.85, 95% CI 0.60-5.72 in the not overweight/obese stratum, and 8.82, 95% CI 0.86-90.33 in the overweight/obese stratum). Second, within *each* level of body satisfaction, we saw fairly high odds across *all* levels of Internet use, even 1-2 hours (eg, for the outcome of very dissatisfied/dissatisfied, comparing 1-2 hours of Internet use vs none, the AOR was 1.48, 95% CI 0.71-3.11 in the original model, 1.85, 95% CI 0.81-4.21 in the not overweight/obese stratum, and 6.49, 95% CI 1.06-39.66 in the overweight/obese stratum). The emergence of strong odds for dissatisfaction with body image across all levels of Internet use are consistent with what one might hypothesize for those who are overweight/obese. Thus, there is some evidence of an interaction effect that should be investigated in future studies, but this dataset is too underpowered to report on these findings as primary conclusions.

Limitations

First, as mentioned, the measure of Internet use available in the CCHS was broad and lacked specificity, referring to time spent on a computer including playing computer games as well as surfing the “World Wide Web.” While computers are used for multiple purposes, many users spend their time online [22], and research suggests the most highly used online platform among girls and young women is Facebook [18], as well as magazine, celebrity, and beauty websites [11,18,19,21]. Further, a recent content analysis of video games found that these types of platforms perpetuate thin representations of female bodies [35]. Further, the exclusion of mobile use to assess Internet use is a strong limitation of this variable, since the Internet is accessed mainly via mobile devices today. Thus, it is likely that the measure used in this study produced a more conservative estimate of the relationship between Internet use and body dissatisfaction. This is a limitation of using population-based data administered by national statistics agencies. While we lack an in-depth validated measure of exposure, the study is strengthened by the use of a large, national, population-based dataset that provides exploration of relationships at the population-based level to inform the next studies. Future research, including forthcoming CCHS cycles, should contain more precise and nuanced measures of Internet use, including social media and mobile use.

Second, self-reporting of body dissatisfaction and Internet use to study interviewers may have been influenced by intentional or nonintentional social desirability bias, and in the case of Internet use, memory distortions. For instance, participants may have overreported their satisfaction level with their body. An underestimation of Internet use is also plausible, especially since figures suggest that many people under 30 are connected almost constantly or at least multiple times a day [36]. Because of this, measuring Internet use up to 20 hours a week could have created a ceiling effect, and interesting relationships between Internet use and body dissatisfaction may have been obscured through collapsing into a single category those people who use the Internet for 20 hours and those who use it almost constantly throughout their day. Both of these occurrences may have attenuated effects. Future research on this topic should consider using technology to track Internet use in real time.

Third, we derived the data from a cross-sectional survey, and one possible theoretical explanation for the findings is that women who start off with poor body satisfaction may be more likely to increase their Internet use. Prospective cohort studies that evaluate this relationship over time are needed.

Fourth, girls and young women are not a homogeneous group. They vary in age, sex (biological), gender (social), ethnicity, income, education, resiliency, and many other social identities and cultural factors (including Internet exposure), all of which can intersect in meaningful ways to influence the development of body dissatisfaction. Our primary focus was to measure the effect of Internet use (ie, we were not concerned with identifying the range of possible predictors). Future research should examine intersectional differences between multiple social categories. The strengths of this study are its large sample size and that it is based on a national population representative of 940,786 Canadian females aged 12-29 years.

Conclusion

Given the number of poor health outcomes associated with body dissatisfaction [7-10], and the significant rise in Internet and social media use [22-24] that allows for unprecedented appearance comparisons [37], public health efforts are needed to support Canadian girls and young women to achieve and maintain a positive body image in today’s digital age. Evidence-based guidelines recommending healthier Internet use could be offered to empower girls and young women with the knowledge and skills to develop an authentic, healthy identity while engaging online. As activists and experts highlight [16,34], this may include limiting daily screen time, abstaining from using social media for 1 to 3 days, paying attention to the kind of content consumed and the impact it may be having on how we feel about ourselves, unfollowing people or sites that cause harmful self-evaluations, and, conversely, following those who promote health and well-being for ourselves. Further, public health education campaigns in schools and online, such as the “More to Her” media literacy curriculum by Raw Beauty Talks [16] or the 8-week Body Image Resilience Program by Beauty Redefined [34], are needed to promote confidence, self-love, and appreciation of oneself beyond what appears on the outside for all girls and women, in all of their diversity. Lastly, action is needed at policy levels to encourage advertisers and corporations to limit the editing of images and promote diverse and realistic representations of women. With Internet use only destined to grow, the time is now to reclaim definitions of female beauty and worth.

Acknowledgments

We would like to thank all of the girls and young women who participated in the Canadian Community Health Survey (CCHS) 2011-2012. We also thank Dr Mieke Koehoorn, Professor at the University of British Columbia, for her critical feedback throughout this research project that was conducted as part of graduate coursework. Her guidance and support, which continued long after coursework was completed, is greatly appreciated. Lastly, we thank the reviewers for their detailed comments and suggestions, which strengthened the final manuscript. Allison Carter received a Doctoral Research Scholarship from the Canadian Institutes of Health Research (CIHR) and the Canadian Association for HIV Research (CAHR). Angela Kaida received salary support through a Tier 2 Canada Research Chair in Global Perspectives on HIV and Sexual and Reproductive Health.

Conflicts of Interest

None declared.

References

1. Cash T. A “negative body image”: evaluating epidemiological evidence. In: Cash T, Pruzinsky TF, editors. Body Image: A Handbook of Theory, Research, and Clinical Practice. New York, NY: Guilford Press; 2002:269-276.
2. Al Sabbah H, Vereecken CA, Elgar FJ, Nansel T, Aasvee K, Abdeen Z, et al. Body weight dissatisfaction and communication with parents among adolescents in 24 countries: international cross-sectional survey. BMC Public Health 2009 Feb 06;9:52-62 [FREE Full text] [doi: 10.1186/1471-2458-9-52] [Medline: 19200369]
3. Dion J, Blackburn M, Auclair J, Laberge L, Veillette S, Gaudreault M, et al. Development and aetiology of body dissatisfaction in adolescent boys and girls. Int J Adolesc Youth 2015 Apr 03;20(2):151-166 [FREE Full text] [doi: 10.1080/02673843.2014.985320] [Medline: 25931646]
4. Pinkasavage E, Arigo D, Schumacher LM. Social comparison, negative body image, and disordered eating behavior: the moderating role of coping style. Eat Behav 2015 Jan;16:72-77. [doi: 10.1016/j.eatbeh.2014.10.014] [Medline: 25464070]
5. Strahan EJ, Wilson AE, Cressman KE, Buote VM. Comparing to perfection: How cultural norms for appearance affect social comparisons and self-image. Body Image 2006 Sep;3(3):211-227. [doi: 10.1016/j.bodyim.2006.07.004] [Medline: 18089224]
6. Cattarin J, Thompson J, Thomas C, Williams R. Body image, mood, and televised images of attractiveness: the role of social comparison. J Soc Clin Psychol 2000;19(2):220-239.
7. van den Berg PA, Mond J, Eisenberg M, Ackard D, Neumark-Sztainer D. The link between body dissatisfaction and self-esteem in adolescents: similarities across gender, age, weight status, race/ethnicity, and socioeconomic status. J Adolesc Health 2010 Sep;47(3):290-296 [FREE Full text] [doi: 10.1016/j.jadohealth.2010.02.004] [Medline: 20708569]
8. Brausch AM, Gutierrez PM. The role of body image and disordered eating as risk factors for depression and suicidal ideation in adolescents. Suicide Life Threat Behav 2009 Feb;39(1):58-71. [doi: 10.1521/suli.2009.39.1.58] [Medline: 19298151]
9. Ferreiro F, Seoane G, Senra C. A prospective study of risk factors for the development of depression and disordered eating in adolescents. J Clin Child Adolesc Psychol 2011;40(3):500-505. [doi: 10.1080/15374416.2011.563465] [Medline: 21534061]
10. Blashill AJ, Safren SA. Body dissatisfaction and condom use self-efficacy: a meta-analysis. Body Image 2015 Jan;12:73-77 [FREE Full text] [doi: 10.1016/j.bodyim.2014.10.002] [Medline: 25462884]
11. Bedford JL, Johnson CS. Societal influences on body image dissatisfaction in younger and older women. J Women Aging 2006;18(1):41-55. [doi: 10.1300/J074v18n01_04] [Medline: 16635949]
12. van den Berg PA, Paxton SJ, Keery H, Wall M, Guo J, Neumark-Sztainer D. Body dissatisfaction and body comparison with media images in males and females. Body Image 2007 Sep;4(3):257-268. [doi: 10.1016/j.bodyim.2007.04.003] [Medline: 18089272]
13. Grabe S, Ward LM, Hyde JS. The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. Psychol Bull 2008 May;134(3):460-476. [doi: 10.1037/0033-2909.134.3.460] [Medline: 18444705]
14. López-Guimerà G, Levine M, Sánchez-Carracedo D, Fauquet J. Influence of mass media on body image and eating disordered attitudes and behaviors in females: a review of effects and processes. J Media Psychol 2010;13(4):387-416.
15. Dove. The Dove Campaign for Real Beauty. London, UK: Unilever; 2016. URL: http://www.dove.us/Social-Mission/campaign-for-real-beauty.aspx [accessed 2016-02-11] [WebCite Cache ID 6fEBdls8D]
16. RAW. Raw Beauty Talks. 2016. URL: http://www.rawbeautytalks.com/ [accessed 2016-02-11] [WebCite Cache ID 6fEC3Sv4j]
17. Qiu L, Lin H, Leung AK, Tov W. Putting their best foot forward: emotional disclosure on Facebook. Cyberpsychol Behav Soc Netw 2012 Oct;15(10):569-572. [doi: 10.1089/cyber.2012.0200] [Medline: 22924675]
18. Tiggemann M, Slater A. NetGirls: the Internet, Facebook, and body image concern in adolescent girls. Int J Eat Disord 2013 Sep;46(6):630-633. [doi: 10.1002/eat.22141] [Medline: 23712456]
19. Tiggemann M, Slater A. NetTweens: the Internet and body image concerns in preteenage girls. J Early Adolesc 2013 Sep 05;34(5):606-620. [doi: 10.1177/0272431613501083]

20. Stronge S, Greaves LM, Milojev P, West-Newman T, Barlow FK, Sibley CG. Facebook is linked to body dissatisfaction: comparing users and non-users. Sex Roles 2015 Aug 5;73(5-6):200-213. [doi: 10.1007/s11199-015-0517-6]
21. Tiggemann M, Miller J. The Internet and adolescent girls' weight satisfaction and drive for thinness. Sex Roles 2010 Apr 28;63(1-2):79-90. [doi: 10.1007/s11199-010-9789-z]
22. CBC News. 2015. Desktop Internet use by Canadians highest in world, comScore says URL: http://www.cbc.ca/news/business/desktop-internet-use-by-canadians-highest-in-world-comscore-says-1.3012666 [accessed 2016-02-11] [WebCite Cache ID 6fEC7h9US]
23. Internet World Stats. 2015. Canada Internet usage, broadband and telecommunications reports URL: http://www.internetworldstats.com/am/ca.htm [accessed 2016-02-11] [WebCite Cache ID 6fEC9c7zm]
24. Bennett S. Social Media Trends: 59% use Facebook in Canada (LinkedIn: 30%, Twitter: 25%, Instagram: 16%). 2015. URL: http://www.adweek.com/socialtimes/canada-social-media-study/614360 [accessed 2016-02-11] [WebCite Cache ID 6fECGwQ0R]
25. Statistics Canada. Canadian Community Health Survey (CCHS) Annual Component: 2011-2012 Microdata File User Guide. Ottawa, ON: Statistics Canada; 2012.
26. Meerkerk G, Van Den Eijnden R, Vermulst AA, Garretsen HFL. The Compulsive Internet Use Scale (CIUS): some psychometric properties. Cyberpsychol Behav 2009 Feb;12(1):1-6. [doi: 10.1089/cpb.2008.0181] [Medline: 19072079]
27. Gucciardi E, Wang SC, Badiani T, Stewart DE. Beyond adolescence exploring Canadian women and men's perception of overweight. Womens Health Issues 2007;17(6):374-382. [doi: 10.1016/j.whi.2007.05.007] [Medline: 17951073]
28. UCLA Institute for Digital Research and Education. SAS data analysis examples: logit regression. 2015. URL: http://www.ats.ucla.edu/stat/sas/dae/logit.htm [accessed 2016-02-11] [WebCite Cache ID 6fECIh3yp]
29. UCLA Institute for Digital Research and Education. SAS data analysis examples: mulitnomial logistic regression. 2015. URL: http://www.ats.ucla.edu/stat/sas/dae/mlogit.htm, [accessed 2016-02-11] [WebCite Cache ID 6fECKRmaI]
30. Bozdogan H. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika 1987 Sep;52(3):345-370. [doi: 10.1007/BF02294361]
31. Brisebois F, Thivierge S. The weighting strategy of the Canadian Community Health Survey. 2001 Presented at: American Statistical Association Meeting, Survey Research Methods Section; August 5-9, 2001; Atlanta, GA, USA.
32. Rothman K. Epidemiology: An Introduction. 2nd edition. Oxford, UK: Oxford University Press; 2012.
33. Austin SB, Haines J, Veugelers PJ. Body satisfaction and body weight: gender differences and sociodemographic determinants. BMC Public Health 2009 Aug 27;9:313-332 [FREE Full text] [doi: 10.1186/1471-2458-9-313] [Medline: 19712443]
34. Kite L, Kite L. Beauty Redefined. Logan, UT: Kite Media; 2014. URL: http://www.beautyredefined.net/ [accessed 2016-02-11] [WebCite Cache ID 6fECNuu6I]
35. Martins N, Williams DC, Harrison K, Ratan RA. A content analysis of female body imagery in video games. Sex Roles 2009 Jul 10;61(11-12):824-836. [doi: 10.1007/s11199-009-9682-9]
36. Perrin A. One-fifth of Americans report going online 'almost constantly'. Washington, DC: Pew Research Centre; 2015 Dec 08. URL: http://www.pewresearch.org/fact-tank/2015/12/08/one-fifth-of-americans-report-going-online-almost-constantly/ [accessed 2016-02-11] [WebCite Cache ID 6fECRutl2]
37. Wasylkiw L, Williamson ME. Actual reports and perceptions of body image concerns of young women and their friends. Sex Roles 2012 Nov 6;68(3):239-251. [doi: 10.1007/s11199-012-0227-2] [Medline: 26811110]

Abbreviations

AOR: adjusted odds ratio
BMI: body mass index
CCHS: Canadian Community Health Survey
OR: odds ratio

Edited by G Eysenbach; submitted 11.02.16; peer-reviewed by S Stronge, J Guiterrez, J Chaplin, DM Liou; comments to author 19.04.16; revised version received 17.05.16; accepted 21.01.17; published 09.02.17.

Please cite as:
Carter A, Forrest JI, Kaida A
Association Between Internet Use and Body Dissatisfaction Among Young Females: Cross-Sectional Analysis of the Canadian Community Health Survey
J Med Internet Res 2017;19(2):e39
URL: http://www.jmir.org/2017/2/e39/
doi:10.2196/jmir.5636
PMID:28183688

Original Paper

Use of Information and Communication Technologies Among Older People With and Without Frailty: A Population-Based Survey

Niina Susanna Keränen[1,2,3], MSc, MD; Maarit Kangas[1,3], PhD; Milla Immonen[4,5], MSc; Heidi Similä[4,5], MSc; Heidi Enwald[3,6], PhD; Raija Korpelainen[3,4,7], PhD; Timo Jämsä[1,3,8], PhD

[1]Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Oulu, Finland
[2]Infotech Oulu, Oulu, Finland
[3]Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland
[4]Center for Life Course Health Research, University of Oulu, Oulu, Finland
[5]VTT Technical Research Centre of Finland Ltd, Oulu, Finland
[6]Information Studies, Faculty of Humanities, University of Oulu, Oulu, Finland
[7]Department of Sports and Exercise Medicine, Oulu Deaconess Institute, Oulu, Finland
[8]Department of Diagnostic Radiology, Oulu University Hospital, Oulu, Finland

Corresponding Author:
Niina Susanna Keränen, MSc, MD
Research Unit of Medical Imaging, Physics and Technology
University of Oulu
POB 5000
Oulu, FI-90014
Finland
Phone: 358 294486008
Fax: 358 8344084
Email: niina.s.keranen@oulu.fi

Abstract

Background: Use of information and communication technologies (ICT) among seniors is increasing; however, studies on the use of ICT by seniors at the highest risk of health impairment are lacking. Frail and prefrail seniors are a group that would likely benefit from preventive nutrition and exercise interventions, both of which can take advantage of ICT.

Objective: The objective of the study was to quantify the differences in ICT use, attitudes, and reasons for nonuse among physically frail, prefrail, and nonfrail home-dwelling seniors.

Methods: This was a population-based questionnaire study on people aged 65-98 years living in Northern Finland. A total of 794 eligible individuals responded out of a contacted random sample of 1500.

Results: In this study, 29.8% (237/794) of the respondents were classified as frail or prefrail. The ICT use of frail persons was lower than that of the nonfrail ones. In multivariable logistic regression analysis, age and education level were associated with both the use of Internet and advanced mobile ICT such as smartphones or tablets. Controlling for age and education, frailty or prefrailty was independently related to the nonuse of advanced mobile ICT (odds ratio, OR=0.61, *P*=.01), and frailty with use of the Internet (OR=0.45, *P*=.03). The frail or prefrail ICT nonusers also held the most negative opinions on the usefulness or usability of mobile ICT. When opinion variables were included in the model, frailty status remained a significant predictor of ICT use.

Conclusions: Physical frailty status is associated with older peoples' ICT use independent of age, education, and opinions on ICT use. This should be taken into consideration when designing preventive and assistive technologies and interventions for older people at risk of health impairment.

(J Med Internet Res 2017;19(2):e29) doi:10.2196/jmir.5507

KEYWORDS

frail elderly; Internet; smartphone; population characteristics; logistic regression; aged

Introduction

The growing role of information and communication technologies (ICT) in our daily lives has led to concerns of increasing inequality between those who can and those who cannot take advantage of new technologies. The divide has been observed not only between younger and older people [1-3], but also within subgroups of older adults [2,4]. This is not only due to lack of access (first-level digital divide) but also due to lack of use (second-level digital divide) [4]. Advanced age, low education, low income, and disability have been shown to predict low Internet use among seniors [2,4-6].

ICT use can also be beneficial for older adults. Use of ICT for direct interaction with people or indirectly for other tasks has also been shown to contribute to the well-being and quality of life of older persons [7] and has been suggested to be associated with social engagement [7,8]. Moreover, ICT use offers a cognitively and intellectually challenging activity that can empower individuals [9]. Computers and Internet also provide many services that support autonomy in old age by facilitating the execution of many routine tasks through e-services (eg, banking, shopping, and communication with social and health services). Older adults who possess better cognitive skills are much better positioned to benefit from the Web-based services available [10].

The question remains, however, how greatly the digital divide among older adults hampers the potential reach of beneficial e-services such as eHealth within the groups that would most benefit from them. Previous studies on the ICT use of older adults with impaired health have used rather crude markers such as existing disability or a medical condition [2,11,12], or self-rated health [12,13]. In all these studies, the metric of poor health has predicted less ICT use, but the metrics of health may not highlight the ICT use and needs of home-dwelling older people most at risk of health impairment, hospitalization, and mortality.

Frailty, as a medical concept, is the loosely defined collection of physiological changes that results in an increased risk of adverse effects in response to stressors. Frail persons are at an elevated risk of falls, declining mobility or disability in daily activities, hospitalization, and death. For a frail person, any minor infection, trauma, or other events can cause a major change in health status and result in loss of independency [14,15]. Demographic factors that have been shown to be associated with frailty largely overlap with factors associated with low ICT use, including advanced age [14,16,17], low education [14,16,17], low income [14,17], and female gender [14,16]. Estimations of the prevalence of frailty vary, but are generally in the region of 10% in the community-dwelling older population [18].

In addition to the general benefits of ICT use for seniors, it could be utilized in various ways in the care of frail seniors in particular. Nutrition such as caloric and protein support [19] and physical activity such as personalized exercise [20,21] have been shown to have an effect on the progression of frailty. These are also potential targets for ICT-assisted interventions [22-25]. Frail older people are also active users of health and social services, yet many of them are deemed by frail users as inflexible, not catering to their needs, or impeded by the person's limited mobility [26]. Frail individuals would benefit from individualized home-delivered health programs [27], which could also be digitally delivered. Mobile devices might also be used for the evaluation of frailty [28]. Additionally, ICT can be used for remote fall detection and rehabilitation among frail seniors [29,30].

In this population-based study, we aimed to compare the frail, prefrail, and nonfrail senior citizens regarding (1) ICT use, defined as use of Internet or advanced mobile devices (smartphones and tablet computers) and (2) attitudes toward the use of advanced mobile devices. We hypothesized that there would be lower ICT use among the frail seniors, and that the difference in ICT use would not be explained by age and education alone. Our aim was to produce new knowledge that can be utilized when developing effective interventions for persons at risk of impaired health and loss of independence.

Methods

The GASEL Survey

The study was conducted from November 2014 to January 2015 in Oulu. Oulu is Finland's fifth largest city with a population of approximately 200,000, and 14% of the population is aged 65 years or more [31]. The study questionnaire was piloted among 11 volunteer seniors, and minor usability-related changes were made based on user feedback. A random sample of 1500 people living in Northern Finland was obtained from the Population Register Centre of Finland. The sampling criteria were as follows: (1) born during January 1, 1900 to December 31, 1949 (aged 65-114 years by the end of 2014); (2) spoke Finnish as a native language; and (3) had a permanent living address in Oulu, Finland, in November 2014. A reminder and another copy of the questionnaire were sent to the nonresponders 4 weeks after the first survey.

The study protocol was approved by the Ethics Committee of Human Sciences in the University of Oulu (statement 6/2014). Questions on ICT use, health, lifestyle, and sociodemographic factors were posed as described in the following sections. The study was part of the GASEL (Gamified Services for Elderly) project.

Assessment of Frailty

Operational indicators of frailty are usually based on weight loss, fatigue, slowness, low physical activity, number of illnesses, and physical attributes such as low grip strength, inability to rise from a chair, or walk upstairs [19,32]. However, no single clear operational definition has been established [19].

To identify the frail study participants, we used a modified version of the 3-item SOF (Study of Osteoporotic Fractures) index, validated for the prediction of falls, disability, fracture, and mortality [32,33]. The modified items and rationale for modifications has been presented in Table 1. According to the responses, the responders were categorized as nonfrail (0 indicators), prefrail (1 indicator), or frail (2-3 indicators).

Table 1. The modified frailty index.

SOF[a] index [13]	Modified item	Reason for modification
Weight loss of 5% in 3-4 years	At least 1 kg weight loss in 3 months	Easier to remember. Part of other scales used in the GASEL study.
Inability to rise from a chair 5 times without using arms	Inability to rise from a chair without using arms, as identified by “Can you rise from a chair independently and without using hands?” answered on a 5-point scale.	Risk of falling if the participants do the test unattended at home.
Poor energy as identified by “No” to “Do you feel full of energy?”	Poor energy as identified by “Which of the following best describes how energetic you have felt in the last month?” and response of “I feel moderately tired, exhausted or weak” or worse.	Fitting in with other scales. Adapted from the 15D instrument [34].

[a]SOF: Study of Osteoporotic Fractures.

Use of the Internet

Internet use in the last 3 months was assessed with the question “How often on average did you use the Internet in the last 3 months?” with 3 frequency response options, and 1 “Not at all.” Internet access at home was assessed with the question “Do you or anyone in your household have access to the Internet at home? (by any device)” with the options “Yes,” “No,” and “I don’t know.” These questions are slightly modified for brevity from the ones used in Eurostat surveys [35], where Internet use or nonuse is a separate question.

Use of Specific Devices Including Smartphones and Tablets

Use of specific devices was asked with the question “In the last 12 months have you used following devices” with the response options “Yes, without major difficulties,” “Yes, with difficulties,” “No,” and “I don't know.” The value of differentiating between use with difficulty and use without difficulty has been previously shown in the development of the everyday technology use questionnaire [36].

Smartphone use was asked with the question “Have you used a mobile phone with a touch display?” Although not all smartphones are touch phones or vice versa, this simplification made responding easier for the less technologically inclined.

Technology acceptance was asked based on the constructs of the Unified Theory of Acceptance and Use of Technology (UTAUT) [37]. Technology acceptance models strive to determine the variables, and the interactions of said variables, that predict the adoption of an available technology by an individual. UTAUT is part of a continuum of popular models that also includes the Technology Acceptance Model (TAM) [38], TAM2 [39], and UTAUT2 [40]. These models have been shown to be valid in a large number of contexts, countries, and population groups, and among older age groups [41].

UTAUT includes 3 constructs that predict behavioral intention to use the specific technology. Performance expectancy (PE) is defined as the degree to which technology provides benefits to users in performing certain activities; effort expectancy (EE) refers to the degree of ease associated with the use of technology; and social influence (SI) refers to the extent to which users perceive that important others (eg, family and friends) believe they should use a particular technology [40]. These determinants are moderated by gender, age, experience, and voluntariness of use. The determinants are usually assessed with a standard set of questions. An additional construct, facilitating conditions, refers to the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system and affects the transition from intention to use [37].

We used UTAUT as a list of factors affecting technology adoption of mobile technologies such as smartphones or tablets. Facilitating conditions could not be assessed because of the many possible types and the length of the questionnaire. Outside of the questions rooted in the UTAUT framework, the questionnaire was supplemented with a general interest question, a question regarding cost, and a question about privacy concerns. Only 1 question per determinant was included and answered with “Yes,” “No,” or “I don’t know.” For regression analyses, “No or I don’t know” was used as the reference category [42].

Socioeconomic Characteristics

The type of habitation of the participants was assessed with the question “Which of the following best describes your current form of residence?,” with 3 options corresponding to single- and multistory buildings with or without elevator, and one “assisted living building, rehabilitation center, or nursing home.”

Economic situation was assessed with the question “Is your household income sufficient for your needs?” with the options of “Yes, very well”; “Yes, reasonably well”; “Barely”; and “No (I am using up my savings or being supported by my friends and family).” The latter 2 were classified in the analyses as financial concerns.

Sensory problems were assessed with the dichotomous questions “Do you have difficulties in close activities (eg, handcrafts, reading) due to poor vision?” and “Do you have difficulties hearing (eg, using the phone)?” with the guidance to answer according to experiences in daily life, using the assistive devices usually used.

Mobility outside of the house was assessed with a combination of 2 questions. The Mini Nutritional Assessment (MNA) embedded within the survey included the question “Which of the following best describes your current mobility?” The response options were “I am bed or chair bound”; “I am able to get out of bed/chair, but do not go out”; and “I am able to move outside the house.” This was supplemented with the question “Do you use mobility aids?” with a list of various mobility aids and the multiple options “Indoors,” “Outdoors,” and “Not at all.” The first 2 MNA response options were considered to indicate indoors only, with the exception of a "Bed or chair bound" person who used mobility aids outdoors.

Loneliness was assessed with the 6-item Gierveld scale, which scores emotional and social loneliness on a scale of 0-6 [43]. We used the questions and the 5-point answer scale for responses previously translated into Finnish for the 11-item long version [44], which includes the 6-item scale as a subset.

Medical conditions were asked with a dichotomous answer to a list of possible diagnosed conditions. Dementia was asked separately as "mild dementia" or "moderate or severe dementia." In cases where a respondent had answered at least one medical condition question, it was assumed blank answers correspond to "No."

Education was asked on a 4-level scale, with the options "Less than primary education"; "Primary education"; "High school or college"; and "University degree or similar." High school or above was considered higher education.

Participation on organized activities outside of the home environment was asked with the question "Do you participate in the activities of a club, union, society, hobby group or spiritual or religious society (eg, sports group, political party, choir, congregation)?" with the answer options "Yes, frequently," "Yes, occasionally," and "No."

Statistical Analysis

The data of each respondent were included in this study if the respondent (1) answered to all the frailty index questions; or scored 2 points in the index (answer to the additional question would not have changed the classification); (2) did not have moderate or severe dementia; and (3) lived in an environment that was not an assisted living facility or nursing home.

Kendall tau was used for crosstabs with ordinal variables and all 3 responder groups. Chi square tests were used to analyze the statistical significance of the differences in dichotomous frailty categories ("nonfrail" and "frail or prefrail"; or "nonfrail or prefrail" and "frail"). Independent-samples Kruskal–Wallis test was used to evaluate the significance of differences in continuous variables (age, number of daily medications) between the groups.

Univariate associations between explanatory variables and ICT use were first analyzed using cross-tabulation. Factors associated with ICT use in univariate analyses were entered into the multivariable logistic regression analysis. Age was first entered into the model, followed by the significant variables using forward stepwise (likelihood ratio) method as a second block. Two models were built and tested, 1 for each of the 2 different dichotomous classifications of frailty.

All statistical analyses were performed using SPSS Statistics version 22, 64-bit edition (IBM). The results were presented as odds ratios (ORs) with 95% CIs. The statistical significance was set at *P* value <.05.

Results

Socioeconomic Characteristics of the Responder Groups

By January 2015, 918 responses had been received, resulting in a response rate of 61.2%. After checking against inclusion criteria, a total of 794 responses (52.9% of the original sample) were included in the analysis. The responders were approximately 2.5 years younger than nonresponders or those who were excluded (*P*<.001), but there was no statistically significant difference in gender distribution (*P*=.42). Fifty-six (7.1%) were classified as frail, 181 (22.8%) as prefrail, and 557 (70.2%) as nonfrail.

Age (*P*<.001), financial concerns (*P*<.001), number of daily medications (*P*<.001), prevalence of mild dementia (*P*<.001), and trouble with near vision (*P*<.001) or hearing (*P*=.001) were positively associated with frailty level, whereas mobility outside home (*P*<.001) and participation in organized activities such as clubs, societies, political parties, or church (*P*<.001) were inversely associated with frailty level (ordinal analyses where applicable). A summary of these findings is shown in Multimedia Appendix 1. There were also significant differences in the average loneliness scores of the different groups (nonfrail 2.05, prefrail 2.63, frail 3.39; *P*<.001).

Use of the Internet

The use of ICT was different across the frailty groups. Frailty level was significantly and inversely associated with having Internet connection at home (78.6% of nonfrail, 70.2% of prefrail, 46.4% of frail, *P*<.001), Internet use in last 3 months (71.8% of nonfrail, 64.1% of prefrail, 33.9% of frail, *P*<.001), and computer use in the last 12 months (70.0% of nonfrail, 62.4% of prefrail, 30.4% of frail, *P*<.001).

In univariate analyses, age, higher education, prevalence of mild dementia, poor near vision, and dichotomous frailty status were associated with Internet use, whereas gender was not (Multimedia Appendix 2). The results of the multivariable logistic regression with these variables are shown in Table 2.

Among those participants who had used the Internet during the last 12 months (n=556), there were no statistically significant differences in the type of Internet use between the different frailty level groups. The most common types of Internet use were e-services such as bank, social services, taxes, and tickets (86.2%); information such as timetables, health information, or recipes (83.3%); news (81.1%); communication such as email or Skype (70.1%); entertainment such as movies, books, and music (40.6%); and gaming (34.9%). Shopping, hobbies or studying, social media, and following or posting on forums were less frequent uses.

Table 2. Socioeconomic predictors of Internet use in the last 3 months among Finnish seniors aged 65+ years, according to multivariable regression analysis adjusted for age (n=738).

Socioeconomic predictors	Model 1[a]			Model 2[b]		
	Odds ratio	95% CI	*P* value	Odds ratio	95% CI	*P* value
Age (years)	0.87	0.84-0.90	<.001	0.87	0.84-0.90	<.001
High education	5.14	3.28-8.07	<.001	4.97	3.17-7.80	<.001
Poor near vision	0.65	0.45-0.94	.02	0.67	0.47-0.98	.04
Frailty				0.45	0.22-0.91	.03

[a] Nonfrail or prefrail is the reference category.

[b] Frail is the reference category.

Use of Specific Devices Including Smartphones and Tablets

The majority of all responders, including the frail participants, used a regular mobile phone. Smartphones and tablets were used by less than a third of the responders and a small minority of frail persons, and the frail group experienced more difficulties in the use of all items. Only a small minority of responders in any group used a computer or mobile phone designed for seniors or other users with special needs (Multimedia Appendix 3).

In univariate analyses, higher education, prevalence of mild dementia, and dichotomous frailty levels were all associated with use of advanced mobile ICT (smartphones or tablets), but gender or poor near vision was not (Multimedia Appendix 4). The results of multivariable regression analysis with these variables are shown in Table 3. The frail people were not compared with the “prefrail or nonfrail” reference category due to the small number of subjects in the “frail, uses advanced mobile ICT” group (n=7).

When exploring opinions of advanced mobile ICT in ordinal analyses including all 3 response options, we found that nonusers had more negative opinions on performance expectancy (*P*<.001), interest (*P*<.001), effort expectancy (*P*=.002), social influence (*P*<.001), and subjective cost (*P*=.04). Differences in privacy concerns were not statistically significant between users and nonusers (*P*=.50). Not all of these differences can be seen in the univariate regression for dichotomized variables, which is presented in Multimedia Appendix 5. Multivariable regression using the UTAUT-derived variables performance expectancy, effort expectancy, social influence, as well as mild dementia and frailty can be seen in Table 4.

The frail nonusers in particular were more negative than nonfrail nonusers on opinion scales (Figure 1). Among nonusers, ordinal frailty status was significantly correlated with performance expectancy (*P*<.001), effort expectancy (*P*=.002), social influence (*P*=.003), privacy concerns (*P*=.001), and interest (*P*=.01), but not with subjective cost (*P*=.07). Among users, there were no such differences; and frailty status was positively correlated with effort expectancy (*P*=.007). However, the number of frail users of advanced mobile ICT was very low.

Table 3. Socioeconomic predictors of the use of advanced mobile information and communication technologies (ICT) in the last 12 months among Finnish seniors aged 65+ years, according to multivariable regression analysis adjusted for age (n=664).

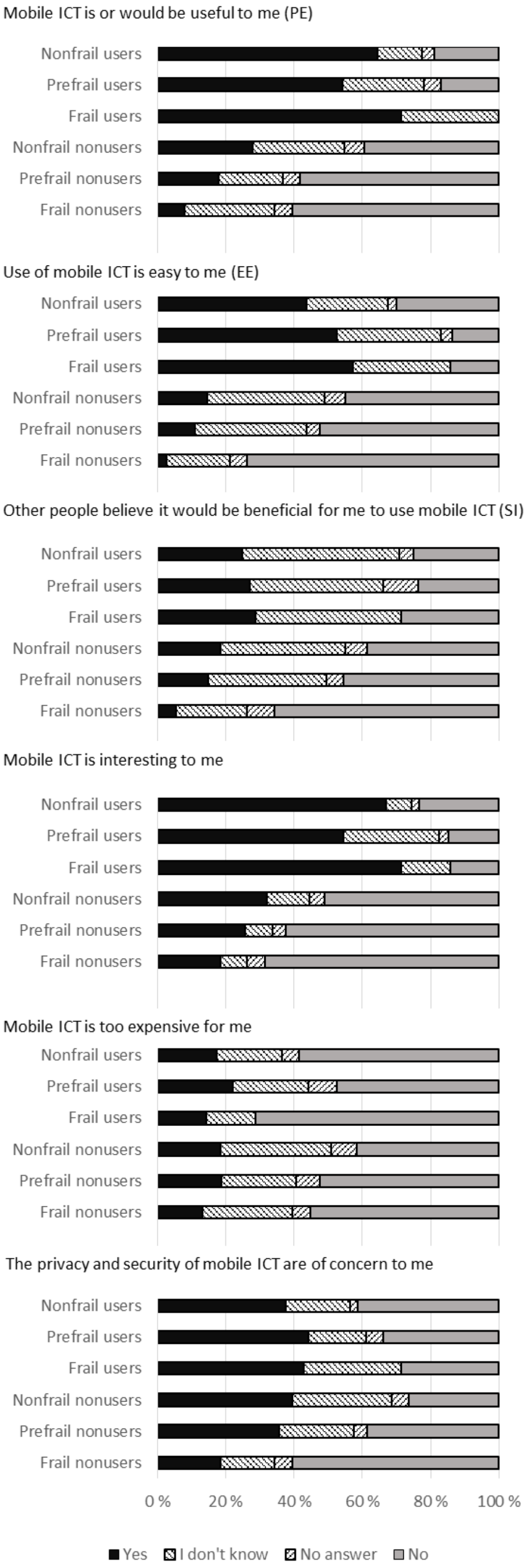
Socioeconomic predictors	Odds ratio	95% CI	*P* value
Age (years)	0.92	0.89-0.95	<.001
High education	2.64	1.89-3.69	<.001
Frailty or prefrailty	0.61	0.42-0.89	.01

Table 4. Attitude predictors of the use of advanced mobile information and communication technologies (ICT) in the last 12 months among Finnish seniors aged 65+ years, according to multivariable regression analysis adjusted for age and education level (n=626).

Attitude predictors	Odds ratio	95% CI	*P* value
Age (years)	0.95	0.92-0.98	.002
High education	1.90	1.31-2.77	.001
Performance expectancy[a]	2.56	1.70-3.98	<.001
Effort expectancy[a]	2.39	1.48-3.86	.001
Frailty or prefrailty	0.60	0.40-0.92	.02

[a] Reference category is those who answered “No” or “I don’t know.”

Figure 1. Potential factors on the adoption of advanced mobile ICT, including questions regarding UTAUT performance expectancy (PE), effort expectancy (EE), and social influence (SI). Users: $n_{nonfrail}$=227; $n_{prefrail}$=57; n_{frail}=7. Nonusers: $n_{nonfrail}$=254; $n_{prefrail}$=101; n_{frail}=38.



Discussion

Principal Findings

Frail seniors represent an important target group for ICT-assisted interventions for the elderly. In this study, we, for the first time, compared the ICT use between frail and nonfrail older people at the population level. The observed lesser ICT use among the frail or prefrail people was not explained by age, education, memory disorders, visual impairments, or opinions on performance or effort to use.

Frailty and low ICT use share risk factors. In this study, the association of Internet use and frailty was only partially explained by the different demographics. The types of Internet use between the frail and nonfrail people were similar. The association of lower use of advanced mobile ICT and frailty was not explained by demographics or attitudes toward the technology.

In general, the frail nonusers were more negative toward advanced mobile ICT than the nonfrail nonusers; there was no such difference in attitudes among users. While the frail respondents reported more difficulties with making ends meet financially, cost was not rated as an obstacle to ICT use among frail nonusers.

Comparison With Prior Work

Our findings are in agreement with previous research results on the prevalence of frailty and general ICT use of the senior population. The prevalence of frailty in this study (7.1%) is similar to previous estimates of 8.6-12.4% of community-dwelling people aged more than 65 years in northern Europe [16]. Here, 67.4% of older people had used the Internet in the last 3 months, and 74.4% had Internet at home. According to Eurostat, 68% of Finnish individuals aged 65-74 years have used the Internet in the last 3 months [3] (2014), and 63% of the age group have Internet access at home [45] (2013). Furthermore, our findings on the demographics of frail persons in Oulu correspond to the previously known higher age, lower education, and lower income than the general senior population. No gender difference was observed. Based on these statistics, we can consider the sample to be representative.

Multiple previous studies have looked into the ICT use of seniors with impaired health, but using different definitions of health. The presence of disability has been shown to be associated with lower Internet use, having fast Internet access at home, less ownership of major digital devices, and less use of email or SMS messages [2,12,46]. Presence of medical conditions [46,47] and low self-rated health [12,13,46] are associated with lower Internet, email, or SMS use. As an addition to these works, frailty is an interesting health classifier. Disability may indicate a functional impairment that may impose restrictions on ICT use while affecting the likelihood of hospitalization or mortality to a much lesser degree, and our findings were not explained by the prevalence of visual impairment. Medical conditions are common among seniors, and approximately half have at least two comorbid conditions [14]; as such it does not alone identify the most high-risk population, whereas the frail or prefrail people represent approximately 10-20% of the age group. Self-rated health has been shown to reflect the risk of mortality and functional impairment [48], but also personality type [49]. It is possible that self-perceived poor health and lack of ICT use share common personality features and general self-perception that are not directly related to health. Physical frailty is also often associated with psychosocial changes [50], such as diminished perception of personal growth and environmental mastery [51], and social isolation and depression [52]; however, metrics of frailty do not rely on the subject’s own evaluation alone.

The original TAM model [38] preceding UTAUT used the constructs “perceived usefulness” and “perceived ease-of-use,” which roughly correspond to PE and EE in UTAUT. These factors have previously been shown to be of chief importance for the technology acceptance of older people [41], and were the most significant predictors of use in our study population as well. Other studies on the nonuse of the Internet by seniors have shown that lack of interest, skepticism of usefulness, and doubt of one’s own skills to use are common attitudes among older ICT nonusers [2,6,53]. We found these same attitudes with regard to advanced mobile ICT. All these negative attitudes were emphasized among frail nonusers, suggesting that enticing this population to use ICT-based services will be particularly difficult.

Our findings offer further evidence that the correlation between low ICT use and health impairment reflects health status and risk, and not simply disability, self-perception, or the personality of the respondent. While older people at large are increasingly using computers, mobile technologies, and the Internet, it is still difficult to reach those most in the need of health care support through such means. The observed similar attitudes among users of mobile ICT regardless of frailty status, as well as the similar uses of the Internet, suggest that there are barriers of entry, rather than different needs. Previous research has divided barriers of ICT adoption with regards to eHealth into (1) physical and psychological attributes; (2) provision; (3) support from others; and (4) economic barriers [53,54]. Based on our study, the difference between frail and nonfrail seniors’ ICT adoption is unlikely to be primarily physical or economic, as neither visual disability nor subjective cost was significant in the multivariable analyses. There are psychosocial features correlated with physical frailty [51,52], which were not included in our models and could discourage the persons from trying new technologies. Possible higher prevalence of social isolation and subsequent lack of recommendations or pressure from peers [4,54] or perceived general helplessness or powerlessness [54] are possible reasons, as are subclinical or undiagnosed cognitive impairments. These should be assessed in further studies targeting health and well-being differences in ICT use.

Limitations

The response rate was high in this study; however, the survey was voluntary. It is likely that the frail are underrepresented in the sample, compared with the general home-dwelling senior population. In particular, the group of frail seniors using advanced mobile ICT was very small, which limited the statistical analyses.

We modified the UTAUT survey for the purposes of this study. The UTAUT model has been built for use with specific use cases rather than broad categories such as “mobile ICT,” and we did not use the standard question set. Thus, the results should only be used for the context of this paper, not as validation of the UTAUT model for this purpose.

Due to the constraints of the survey format and the large number of questions, there were some limitations in the depth of assessment. In particular, controlling for more detailed assessments of cognitive functioning, depressive symptoms, and social network would have been interesting, and should be taken into account in future studies.

Conclusions

This study provides early evidence that community-dwelling nonfrail and frail older people are highly different as ICT users, both in actual use rate and in opinions affecting adoption among the nonusing population. These differences are only partly explained by more advanced age, lower education, and higher prevalence of visual or cognitive impairments. As such the research also highlights the clustering of various psychological and social risk factors of social exclusion to the same physically vulnerable population, and the continuing need of measures to prevent an increasing digital divide in society. Future research is needed to further assess the ICT-related needs, barriers to adoption, and facilitating conditions of this interesting and important population.

Acknowledgments

This study was performed as part of the GASEL project, funded by the Finnish Funding Agency for Innovation (Grant No. 40270/13), BelleGames Oy, BonWell Intelligence Oy, Caritas foundation, Iscom Oy, Mawell Oy, Northern Ostrobothnia Hospital District, Luleå University of Technology, and Oulu Deaconess Institute. The study was also supported by the ASTS joint project by Academy of Finland (Grant No. 270816) and Japan Science and Technology Agency; the iKaaS joint project by the European Union (Grant No. 643262) and Ministry of Internal Affairs and Communication, Japan; Infotech Oulu; and the Finnish Cultural Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of the study participants according to frailty status.

[PDF File (Adobe PDF File), 16KB - jmir_v19i2e29_app1.pdf]

Multimedia Appendix 2

Socioeconomic predictors of Internet use in the last 3 months among Finnish seniors aged 65+ years in univariate regression analysis.

[PDF File (Adobe PDF File), 20KB - jmir_v19i2e29_app2.pdf]

Multimedia Appendix 3

Information and communication technologies (ICT) device use during the last 12 months and difficulties among users across the frailty categories.

[PDF File (Adobe PDF File), 20KB - jmir_v19i2e29_app3.pdf]

Multimedia Appendix 4

Socioeconomic predictors of the use of advanced mobile information and communication technologies (ICT; smartphones or tablets) in the last 12 months among Finnish seniors aged 65+ years in univariate regression analysis.

[PDF File (Adobe PDF File), 20KB - jmir_v19i2e29_app4.pdf]

Multimedia Appendix 5

Attitude predictors of the use of advanced mobile information and communication technologies (ICT) in the last 12 months among Finnish seniors aged 65+ years in univariate regression. “I don’t know” answers counted as negative.

[PDF File (Adobe PDF File), 18KB - jmir_v19i2e29_app5.pdf]

References

1. Smith A. Pew Research Center. Chapter One: A Portrait of Smartphone Ownership URL: http://www.pewinternet.org/2015/04/01/chapter-one-a-portrait-of-smartphone-ownership/ [accessed 2016-12-08] [WebCite Cache ID 6eIteaZ20]
2. Smith A. Pew Research Center. 2014. Older Adults and Technology Use URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2016-12-08] [WebCite Cache ID 6eLzt6ayH]
3. Eurostat. 2014. Individuals-Internet use URL: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ci_ifp_iu&lang=en [accessed 2016-12-08]
4. Friemel TN. The digital divide has grown old: Determinants of a digital divide among seniors. New Media & Society 2014 Jun 12;18(2):313-331. [doi: 10.1177/1461444814538648]
5. Wagner N, Hassanein K, Head M. Computer use by older adults: a multi-disciplinary review. Comput Human Behav 2010 Sep;26(5):870-882. [doi: 10.1016/j.chb.2010.03.029]
6. Statistics Finland. 2013. Vanhojen ikäryhmien tieto- ja viestintätekniikan käyttö [In Finnish] URL: http://www.stat.fi/til/sutivi/2013/sutivi_2013_2013-11-07_kat_002_fi.html [accessed 2016-12-08] [WebCite Cache ID 6eM0CrsMx]
7. Ihm J, Hsieh YP. The implications of information and communication technology use for the social well-being of older adults. Inf Commun Soc 2015 Mar 09;18(10):1123-1138. [doi: 10.1080/1369118X.2015.1019912]
8. Cotten SR, Anderson WA, McCullough BM. Impact of internet use on loneliness and contact with others among older adults: cross-sectional analysis. J Med Internet Res 2013;15(2):e39 [FREE Full text] [doi: 10.2196/jmir.2306] [Medline: 23448864]
9. Amichai-Hamburger Y, McKenna KY, Tal S. E-empowerment: empowerment by the Internet. Comput Human Behav 2008 Sep;24(5):1776-1789. [doi: 10.1016/j.chb.2008.02.002]
10. Freese J, Rivas S, Hargittai E. Cognitive ability and Internet use among older adults. Poetics 2006 Aug;34(4-5):236-249. [doi: 10.1016/j.poetic.2006.05.008]
11. Selwyn N, Gorard S, Furlong J, Madden L. Older adults' use of information and communications technology in everyday life. Ageing Soc 2003 Sep 2;23(05):561-582. [doi: 10.1017/S0144686X03001302]
12. Carpenter BD, Buday S. Computer use among older adults in a naturally occurring retirement community. Comput Human Behav 2007 Nov;23(6):3012-3024. [doi: 10.1016/j.chb.2006.08.015]
13. Chen Y, Persson A. Internet use among young and older adults: relation to psychological well-being. Educ Gerontol 2002 Oct;28(9):731-744. [doi: 10.1080/03601270290099921]
14. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [Medline: 11253156]
15. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762 [FREE Full text] [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
16. Santos-Eggimann B, Cuénoud P, Spagnoli J, Junod J. Prevalence of frailty in middle-aged and older community-dwelling Europeans living in 10 countries. J Gerontol A Biol Sci Med Sci 2009 Jun;64(6):675-681 [FREE Full text] [doi: 10.1093/gerona/glp012] [Medline: 19276189]
17. Woods NF, LaCroix AZ, Gray SL, Aragaki A, Cochrane BB, Brunner RL, Women's Health Initiative. Frailty: emergence and consequences in women aged 65 and older in the women's health initiative observational study. J Am Geriatr Soc 2005 Aug;53(8):1321-1330. [doi: 10.1111/j.1532-5415.2005.53405.x] [Medline: 16078957]
18. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012 Aug;60(8):1487-1492. [doi: 10.1111/j.1532-5415.2012.04054.x] [Medline: 22881367]
19. Morley JE, Vellas B, van Kan GA, Anker SD, Bauer JM, Bernabei R, et al. Frailty consensus: a call to action. J Am Med Dir Assoc 2013 Jun;14(6):392-397 [FREE Full text] [doi: 10.1016/j.jamda.2013.03.022] [Medline: 23764209]
20. de Labra C, Guimaraes-Pinheiro C, Maseda A, Lorenzo T, Millán-Calenti JC. Effects of physical exercise interventions in frail older adults: a systematic review of randomized controlled trials. BMC Geriatr 2015;15:154 [FREE Full text] [doi: 10.1186/s12877-015-0155-4] [Medline: 26626157]
21. Aguirre LE, Villareal DT. Physical exercise as therapy for frailty. Nestle Nutr Inst Workshop Ser 2015;83:83-92 [FREE Full text] [doi: 10.1159/000382065] [Medline: 26524568]
22. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
23. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012 Nov 21;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]
24. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
25. Broekhuizen K, de Gelder J, Wijsman CA, Wijsman LW, Westendorp RGJ, Verhagen E, et al. An Internet-based physical activity intervention to improve quality of life of inactive older adults: a randomized controlled trial. J Med Internet Res 2016;18(4):e74 [FREE Full text] [doi: 10.2196/jmir.4335] [Medline: 27122359]

26. Themessl-Huber M, Hubbard G, Munro P. Frail older people's experiences and use of health and social care services. J Nurs Manag 2007 Mar;15(2):222-229. [doi: 10.1111/j.1365-2834.2007.00726.x] [Medline: 17352706]
27. Turner G, Clegg A, British Geriatrics Society, Age UK, Royal College of General Practioners. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age Ageing 2014 Nov;43(6):744-747 [FREE Full text] [doi: 10.1093/ageing/afu138] [Medline: 25336440]
28. Fontecha J, Hervás R, Bravo J, Navarro FJ. A mobile and ubiquitous approach for supporting frailty assessment in elderly people. J Med Internet Res 2013;15(9):e197 [FREE Full text] [doi: 10.2196/jmir.2529] [Medline: 24004497]
29. Johansson T, Wild C. Telerehabilitation in stroke care--a systematic review. J Telemed Telecare 2011;17(1):1-6. [doi: 10.1258/jtt.2010.100105] [Medline: 21097560]
30. Khan F, Amatya B, Kesselring J, Galea MP. Telerehabilitation for persons with multiple sclerosis. a Cochrane review. Eur J Phys Rehabil Med 2015 Jun;51(3):311-325 [FREE Full text] [Medline: 25943744]
31. Pxnet2.stat. 2013. Väestöllisiä tunnuslukuja alueittain [in Finnish] URL: http://pxnet2.stat.fi/PXWeb/pxweb/fi/StatFin/StatFin__vrm__vaerak/048_vaerak_tau_203.px [accessed 2016-12-08] [WebCite Cache ID 6mbJY2jIm]
32. Ensrud KE, Ewing SK, Cawthon PM, Fink HA, Taylor BC, Cauley JA, Osteoporotic Fractures in Men Research Group. A comparison of frailty indexes for the prediction of falls, disability, fractures, and mortality in older men. J Am Geriatr Soc 2009 Mar;57(3):492-498 [FREE Full text] [doi: 10.1111/j.1532-5415.2009.02137.x] [Medline: 19245414]
33. Ensrud KE, Ewing SK, Taylor BC, Fink HA, Cawthon PM, Stone KL, et al. Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women. Arch Intern Med 2008 Feb 25;168(4):382-389. [doi: 10.1001/archinternmed.2007.113] [Medline: 18299493]
34. Sintonen H, Richardson J. Citeseerx.ist. 1994. The 15-D Measure of Health Related Quality of Life: Reliability, Validity and Sensitivity of its Health State Descriptive System URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.3283 [WebCite Cache ID 6eMBYgVTl]
35. Eurostat. 2014. European Union survey on ICT usage in households and by individuals - Eurostat model questionnaire (version 3.4) URL: https://circabc.europa.eu/sd/a/9d890128-8925-4b8a-9dba-3e9281de3759/Questionnaire%20HH%202014_v3.4.pdf [accessed 2016-01-07] [WebCite Cache ID 6eM0LtO8U]
36. Rosenberg L, Nygård L, Kottorp A. Everyday technology use questionnaire: psychometric evaluation of a new assessment of competence in technology use. OTJR: Occupation, Participation and Health 2008 Dec 19;29(2):52-62. [doi: 10.3928/15394492-20090301-05]
37. Venkatesh V, Morris M, Gordon B, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
38. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
39. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
40. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178.
41. Chen K, Chan AHS. A review of technology acceptance by older adults. Gerontechnology 2011;10(1):1-12. [doi: 10.4017/gt.2011.10.01.006.00]
42. Groothuis P, Whitehead J. Does don't know mean no? Analysis of 'don't know' responses in dichotomous choice contingent valuation questions. Appl Econ 2002;34(15):1935-1940.
43. Gierveld JDJ. A 6-Item scale for overall, emotional, and social loneliness: confirmatory tests on survey data. Res Aging 2006 Sep 01;28(5):582-598. [doi: 10.1177/0164027506289723]
44. Hartikka K. Theseus.fi.: Diaconia University of Applied Sciences; 2012. Vanhusten yksinäisyys palvelutaloissa [in Finnish] URL: https://www.theseus.fi/handle/10024/46796 [accessed 2016-12-08] [WebCite Cache ID 6lAvxUx4E]
45. Eurostat. 2013. Broadband and connectivity - individuals URL: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_bde15b_i&lang=en [accessed 2016-12-09]
46. Gell NM, Rosenberg DE, Demiris G, LaCroix AZ, Patel KV. Patterns of technology use among older adults with and without disabilities. Gerontologist 2015 Jun;55(3):412-421. [doi: 10.1093/geront/gnt166] [Medline: 24379019]
47. Romano MF, Sardella MV, Alboni F, Russo L, Mariotti R, Nicastro I, et al. Is the digital divide an obstacle to e-health? An analysis of the situation in Europe and in Italy. Telemed J E Health 2015 Jan;21(1):24-35. [doi: 10.1089/tmj.2014.0010] [Medline: 25495564]
48. Idler EL, Kasl SV. Self-ratings of health: do they also predict change in functional ability? J Gerontol B Psychol Sci Soc Sci 1995 Nov;50(6):S344-S353. [Medline: 7583813]
49. Goodwin R, Engstrom G. Personality and the perception of health in the general population. Psychol Med 2002 Feb;32(2):325-332. [Medline: 11866326]
50. Schulz R, Williamson GM. Psychosocial and behavioral dimensions of physical frailty. J Gerontol 1993 Sep;48 Spec No:39-43. [Medline: 8409239]
51. Andrew MK, Fisk JD, Rockwood K. Psychological well-being in relation to frailty: a frailty identity crisis? Int Psychogeriatr 2012 Aug;24(8):1347-1353. [doi: 10.1017/S1041610212000269] [Medline: 22436131]

52. Mulasso A, Roppolo M, Giannotta F, Rabaglietti E. Associations of frailty and psychosocial factors with autonomy in daily activities: a cross-sectional study in Italian community-dwelling older adults. Clin Interv Aging 2016;11:37-45 [FREE Full text] [doi: 10.2147/CIA.S95162] [Medline: 26811675]
53. LeRouge C, Van Slyke C, Seale D, Wright K. Baby boomers' adoption of consumer health technologies: survey on readiness and barriers. J Med Internet Res 2014;16(9):e200 [FREE Full text] [doi: 10.2196/jmir.3049] [Medline: 25199475]
54. Jones R. Development of a questionnaire and cross-sectional survey of patient eHealth readiness and eHealth inequalities. Med 2 0 2013;2(2):e9 [FREE Full text] [doi: 10.2196/med20.2559] [Medline: 25075244]

Abbreviations

EE: effort expectancy
FAC: facilitating conditions
ICT: information and communication technologies
PE: performance expectancy
SI: social influence
UTAUT: Unified Theory of Acceptance and Use of Technology
SOF: Study of Osteoporotic Fractures

Edited by G Eysenbach; submitted 03.03.16; peer-reviewed by J Fanning, R Jones; comments to author 24.03.16; revised version received 25.05.16; accepted 29.10.16; published 14.02.17.

Please cite as:
Keränen NS, Kangas M, Immonen M, Similä H, Enwald H, Korpelainen R, Jämsä T
Use of Information and Communication Technologies Among Older People With and Without Frailty: A Population-Based Survey
J Med Internet Res 2017;19(2):e29
URL: http://www.jmir.org/2017/2/e29/
doi:10.2196/jmir.5507
PMID:28196791

Original Paper

Internet Use by Parents of Children With Rare Conditions: Findings From a Study on Parents' Web Information Needs

Honor Nicholl[1], RCN, RGN, MEd, PhD; Catherine Tracey[1], RGN, RPN, MSc, MBA, PhD; Thelma Begley[1], RGN, RCN, RNT, BNS (Hons), MSc; Carole King[1], RCN, RGN, RCT, RNT, MSc; Aileen M Lynch[1], BSc (Hons), MSc, PhD

School of Nursing and Midwifery, Trinity College Dublin, the University of Dublin, Dublin, Ireland

Corresponding Author:
Aileen M Lynch, BSc (Hons), MSc, PhD
School of Nursing and Midwifery
Trinity College Dublin, the University of Dublin
24 D'Olier Street
Dublin, D2
Ireland
Phone: 353 18968571
Fax: 353 18963001
Email: amlynch@tcd.ie

Abstract

Background: Parents of children with rare conditions increasingly use the Internet to source information on their child's condition. This study reports on part of a larger study whose overall aim was to identify the Internet use by parents when seeking information on their child's rare condition, with the specific purpose of using the findings to aid in the development of a website specifically designed to meet the parents' needs. It presents findings on why these parents use the Internet, the information and support content they source, and the impact these resources have on their capacity to care for and manage their child's condition.

Objective: To (1) ascertain parents' general Internet usage patterns, (2) identify the nature of the information parents most frequently searched for, and (3) determine the effect the Internet-sourced information had on parents of children with rare conditions.

Methods: Data collection was conducted in 2 parts: Part 1 was a focus group interview (n=8) to inform the development of the questionnaire, and Part 2 was a questionnaire (Web- and paper-based). All respondents (N=128) completed the questionnaire using the Internet.

Results: Parents frequently and habitually used the Internet and social media to gather information on their child's condition. These Web-based resources provide parents with a parent-to-parent support platform that allows them to share their experiences and information with other parents, which, the respondents considered, improved their knowledge and understanding of their child's condition. The respondents also reported that these resources positively impacted on their decision making, care, and management of their child's condition. However, they reported receiving mixed responses when wishing to engage and share with health care professionals their Internet and social media interactions and information outcomes.

Conclusions: This study adds to the emerging body of research on the Internet use by parents of children with rare conditions to source information on their child's condition. The evolving and ever increasing parent-to-parent support systems via social media are impacting on parents' capacity to manage their children. Implications for practice include health care professionals' response to this knowledge and capacity shift, and the significance of these changes when interacting with parents. The key message of this study was that parents of children with rare conditions are habitual users of the Internet to source information about their children's conditions. Social media, especially Facebook, has an increasing role in the lives of these parents for information and support. Parents' interest in information gathering and sharing includes a desire for shared dialogue with health care professionals.

(J Med Internet Res 2017;19(2):e51) doi:10.2196/jmir.5834

KEYWORDS

rare diseases; parents; Internet; consumer health information; social media; professional-patient relations

Introduction

The Internet is recognized as a significant source of health information [1] and may assist parents of children with rare conditions (*rare conditions* includes rare diseases and rare disorders in this paper) to find information, clarify, and understand the information they have been given, or confirm a diagnosis [2,3]. Parents of these children can experience difficulties with aspects of their child's diagnosis and management, as each condition is rare and diverse in symptoms. Traditionally, health care professionals have been the gatekeepers of specialized knowledge about the child's rare condition [4]. In some cases, the rareness of the child's condition means that information is difficult to find, even for health care professionals, thus making diagnosis and management of the child and his or her condition difficult [4,5].

The Internet is changing the way parents of children with rare conditions access information. It enables them access worldwide information that previously was unavailable to them [6,7]. In comparison to more traditional sources, the Internet may provide large volumes of information from a variety of sources and perspectives, for example, emerging treatment options or relevant clinical trials. This information can be accessed as required, and anonymously and confidentially, if desired. In addition, it is cost-effective and convenient and can be accessed without having to leave one's home [8]. However, when there is no monitoring of the quality or reliability of Internet-based information, the resultant information can be questionable. Parents can also be overwhelmed with both the amount of information available as well as the presentation of case studies of children with more complex or advanced presentations of conditions similar to those of their child.

Through the use of smartphones and mobile devices, information searching via social media and search engines is now a daily occurrence for many people [1]. Additionally, social networks can be established with other parents. Nationally and internationally based Facebook groups or pages are now common for rare conditions [7,9]. Furthermore, parents are using the Internet to receive information and emotional support from Web-based groups of parents facing similar challenges, while also critiquing care provision and health care policy [10]. However, information may not be presented in a balanced way via social networks. Thus, parents need to establish how reliable the information they are acquiring from Internet or social media resources is.

The aim of this study was to identify the Internet-sourced information and support content that meets the needs of parents of children with rare conditions for the purposes of developing an Irish website specifically designed to meet those needs. The objectives were to (1) ascertain parents' general Internet usage patterns, (2) identify the nature of the information parents most frequently searched for, and (3) determine the effect the Internet-sourced information had on parents of children with rare conditions.

Methods

This descriptive exploratory study used both qualitative and quantitative designs. Purposive nonprobability sampling was adopted where participant recruitment was facilitated by the Saoirse Foundation, which acted as the study's gatekeeper. The study was advertised for 4 weeks on the gatekeeper's website and in various linked voluntary organizations and social media links (Twitter, Facebook, blogs). The inclusion criterion was parents (18 years or older) of a child with a rare condition who used the Internet to source information about their child's condition. A more detailed information leaflet was provided to the gatekeeper to distribute to parents who wished to know more about the study. Parents who wanted to participate in the study contacted the researcher and were scheduled to either (1) attend a focus group or (2) complete a questionnaire (Multimedia Appendix 1) that was open for 4 weeks.

The questionnaire was modeled on 2 previously published questionnaires by Porter and Edirippulige [11] and Tozzi et al [12], and 8 participants took part in the focus group to further refine the questionnaire. Written consent was sought prior to the commencement of the focus group. Once the focus group was completed, responses were fed back to the group, thus allowing participants verify their discussion, which was subsequently analyzed and categorized into themes (simple thematic analysis) [13].

The questionnaire consisted of 45 questions that were in the multiple-choice format apart from 2 open questions: (1) *the website(s) I most frequently visit is(are)* and (2) *if you were creating a website for parents of children with rare conditions, what is the ONE thing you would like to see on that website?* The resultant responses were analyzed by simple thematic analysis [13]. The questionnaire was divided into 4 sections, and although 128 respondents started the questionnaire, 93 respondents completed the questionnaire in full. Section 1 contained questions to confirm that the respondent was a parent of a child with a rare condition and used the Internet to source information about their child's condition (inclusion criterion). A total of 7 respondents did not meet this inclusion criterion and were disqualified. So, the remaining respondents were only parents (Table 1).

Table 1. Overview of the questionnaire.

Section number	Section title	Number of questions	Number of respondents who commenced the section
1	Sources of information about your child's condition	3	128 (7 disqualified, as they did not fulfill the inclusion criterion)
2	Information about your use of the Internet to find information about your child's condition	23	121
3	Information about your child or children	8 (per child)	96
4	Information about yourself	11	93
	Total	45	93

Although participants had the option of completing Web- or paper-based versions of the questionnaire, all chose the Web-based option on SurveyMonkey (SurveyMonkey Inc). Ethical approval was obtained from the School of Nursing and Midwifery's Research Ethics Committee, Trinity College Dublin, Ireland.

Results

Demographical Details

In Section 4 of the questionnaire, 93 respondents offered information about themselves (Table 2). The majority were mothers (87%, 81/93), aged between 35 and 49 years (67%, 62/93), living in urban areas (74%, 69/93) mostly in the Republic of Ireland (84%, 78/93). Half of the respondents (50%, 46/93) had a University degree, and 43% (40/93) were their child's main carer (Table 2). Respondents who were not employed full time (84%, 78/93) were asked whether they left full-time employment to care for their child. Of these, 65% (51/78) selected *yes*, and the remaining 35% (27/78) selected either *no* or *not applicable*.

Table 2. Parents' demographics (n=93).

Demographics		n (%)
Gender		
	Females	81 (87)
	Males	12 (13)
Age (in years)		
	18-34	22 (24)
	35-49	62 (67)
	50-64	9 (10)
Country of residence		
	Republic of Ireland	78 (84)
	Northern Ireland	6 (7)
	United Kingdom	7 (8)
	United States	2 (2)
Area of residence		
	City	20 (22)
	Town	30 (32)
	Village	19 (20)
	Rural	24 (26)
Highest level of education		
	Secondary school	23 (25)
	Vocational training	24 (26)
	University: Undergraduate degree	25 (27)
	University: Postgraduate degree	21 (23)
Employment status		
	Employed full time	15 (16)
	Employed part time	20 (22)
	Self-employed	3 (3)
	Your child or children's main carer	40 (43)
	Homemaker	12 (13)
	Student	1 (1)
	Unemployed	2 (2)

In total, 117 children were reported on by their parents (Table 3). Furthermore, 78 parents reported on 1 child, 14 parents reported on 2 children, 1 parent reported on 3 children, and 2 parents reported on 4 children. Over half (54.7%, 64/117) the children were male and ranged in age from less than 12 months to 39 years, with the 4- to 7-year-old category being the largest (28.2%, 33/117). Ninety per cent (105/117) of children had a diagnosis, and 65.7% (69/105) received a diagnosis in their first 3 years of life. The majority (78.6%, 92/117) of children had a disability, and of these, a little more than half (53%, 49/92) had a physical and intellectual disability, 30% (28/92) had a physical disability, 10% (9/92) had an intellectual disability, and there was an equal spread (3%, 3/92) across sensory and neurodevelopmental disabilities, respectively. Some children (39.3%, 46/117) did not need assistive equipment. The remaining 71 children mainly needed equipment for moving (41.0%, 48/117), eating (19.7%, 23/117), and breathing (11.1%, 13/117; Table 3).

Table 3. Children’s demographics.

Demographics			n (%)
Gender (n=117)			
		Female	53 (45.3)
		Male	64 (54.7)
Age (n=117)			
		Under 12 months	5 (4.3)
		1-3	24 (20.5)
		4-7	33 (28.2)
		8-12	28 (23.9)
		13-19	15 (12.8)
		20-29	9 (7.7)
		30-39	3 (2.6)
Does child have a diagnosis (n=117)			
		No	12 (10.3)
		Yes	105 (89.7)
	If yes, age at diagnosis (n=105)		
		Under 12 months	36 (34.3)
		1-3	33 (31.4)
		4-7	16 (15.2)
		8-12	15 (14.3)
		13-19	4 (3.8)
		20-29	1 (0.9)
Does the child's condition include a disability (n=117)			
		No	25 (21.4)
		Yes	92 (78.6)
	If yes, what is the nature of the disability (n=92)		
		Physical and intellectual	49 (53)
		Physical	28 (30)
		Intellectual	9 (10)
		Sensory	3 (3)
		Neurodevelopmental	3 (3)
Does your child use equipment for[a] (n=117)			
		None required	46 (39.3)
		Moving	48 (41.0)
		Eating	23 (19.7)
		Breathing	13 (11.1)
		Speech	10 (8.6)
		Sitting	7 (5.9)
		Hearing	6 (5.1)
		Bathing; sleeping	3 (2.6)
		Writing; nonverbal; seeing	2 (1.7)
		Sensory integration	1 (0.9)

[a] Participants could choose as many options as applied to them.

Internet Use

Respondents' Internet usage is summarized in Table 4. The Internet was used predominantly at home (92.9%, 106/114), between 7 pm to midnight (42.9%, 49/114) and searched on a weekly basis (50.0%, 57/114; either *daily*, *several times*, or *once per week*) mostly on a PC or Mac (41.4%, 46/111). When searching the Internet for information on their child's condition, 45.6% (52/114) *sometimes* found relevant information whereas 37.7% (43/114) found relevant information *most of the time*. All respondents (n=112) had an email address and 94.6% (106/112) had a Facebook account. Most (93.1%, 95/102) used general search engines (Google, Yahoo, Bing), and 28.4% (29/102) accessed specialized sites (Table 4). A total of 75 named websites were accessed (Multimedia Appendix 2).

Table 4. Internet use.

Internet usage		n (%)
From where do you most often access the Internet? (n=114)		
	Home	106 (92.9)
	Work	7 (6.1)
	Public library	0 (0)
	Other (Smartphone)	1 (0.9)
What time of the day do you MOST OFTEN use the Internet (n=114)		
	Midnight to 6 am	2 (1.8)
	7 am to midday	7 (6.1)
	1 pm to 6 pm	10 (8.8)
	7 pm to midnight	49 (42.9)
	No pattern	46 (40.4)
How often do you use the Internet to find information about your child’s condition? (n=114)		
	Every day	23 (20.2)
	Once a week	9 (7.9)
	Several times a week	25 (21.9)
	Once a month	10 (8.8)
	Several times a month	13 (11.4)
	Every few months	25 (21.9)
		9 (7.9)
What device do you use MOST OFTEN to access the Internet? (n=111)		
	PC or Mac	46 (41.4)
	Smartphone	42 (37.8)
	Tablet (iPad or similar)	23 (20.7)
When using the Internet, how often are you able to find the information you are looking for? (n=114)		
	Always	2 (1.8)
	Most of the time	43 (37.7)
	Sometimes	52 (45.6)
	Hardly ever	15 (13.2)
	Never	2 (1.8)
Which of the following applies to you?[a] (n=112)		
	Email	112 (100.0)
	Facebook	106 (94.6)
	Twitter	48 (42.9)
	Skype	57 (50.9)
	LinkedIn	31 (27.7)
	MSN or Messenger	23 (20.5)
	Blog	12 (10.7)
	Health-related apps	9 (8.0)
How do you locate information about your child’s condition?[a] (n=102)		
	Search engines (Google, Yahoo, Bing, etc)	95 (93.1)

Internet usage		n (%)
	Specialized websites	29 (28.4)
	Orphanet	7 (6.9)
	Social media	7 (6.9)
	Recommendations from others	4 (3.9)

[a] Participants could choose as many options as applied to them.

Respondents were asked to rank which factors they took into account when choosing a website and could choose as many factors as applied to them. Of the 112 respondents who replied to this question, 65.2% (73/112) respondents ranked *relevant and accurate* as being the most important factor, closely followed by *trustworthiness* (62.5%, 70/112) and the website being *up to date* (60.7%, 68/112). In order of ranking, other factors included the website being *recommended to me by a health professional* (46.4%, 52/112), *easy to understand* (37.5%, 42/112), has *other websites linked within it* (8.0%, 9/112), and has a *nice layout* (5.4%, 6/112).

Respondents (n=101) ranked which information they searched for (1) when their child was first diagnosed or when they first had a concern about their child's health and (2) what they search for at present (Table 5).

Table 5. Information topics sought by parents of children with rare conditions on the Internet (n=101; Participants who answered the questions could choose as many options as applied to them)

Rank	n[a]	Information searched for when first diagnosed	Change in rank	Rank	n[a]	Information searched for now
1	87	My child’s condition or symptoms	↔	1	71	My child’s condition or symptoms
2	77	My child’s diagnosis	↓2	2	70	The management of my child’s condition
3	69	The management of my child’s condition	↑1	3	51	The care of my child’s condition
						Treatments
4	66	Treatments	↑1	4	47	My child’s diagnosis
		Support groups	↓1			
5	64	Genetics	↓3	5	43	Support groups
6	57	The care of my child’s condition	↑3	6	42	Child development
		Child development	↔			
7	46	Organizations and societies	↓2	7	36	Research and innovation
8	40	Medical or health care professionals	↓2	8	35	Genetics
9	33	Research and innovation	↑2	9	33	Educational options
						Organizations and or societies
10	31	Early intervention options	↓7	10	30	Medical or health care professionals
11	30	Physical activities	↓1	11	29	Preventing complications
						Upcoming events or workshops
12	29	Hospitals, hospices, medical centers	↓3	12	28	Physical activities
13	28	Medicines or alternative treatments or therapies	↓1	13	27	Nutrition
		Nutrition	↔			
14	27	Preventing complications	↑3	14	22	Medicines or alternative treatments or therapies
						Financial assistance
15	26	Future pregnancies	↓2	15	21	Hospitals, hospices, medical centers
						State services
16	25	Educational options	↑7	16	17	Managing family dynamics
17	24	State services	↑2	17	12	Early intervention options
		Financial assistance	↑3			Future pregnancies
18	16	Where to get a second opinion	↓2	18	11	Accessing medicines or alternative treatments or therapies on the Internet
19	15	Upcoming events or workshops	↑8	19	8	Vaccinations
20	12	Managing family dynamics	↑4	20	7	Where to get a second opinion
21	11	Accessing medicines or alternative treatments or therapies on the Internet	↑3			
22	10	Vaccinations	↑3			

[a] Represents the number of participants who chose each information topic.

Overall there was not great variation in the ranking of information searched for over the 2 timelines. The first 10 ranked information topics underwent some minor reordering but remained in the top 10 across both timelines (apart from *early intervention options*), and *my child’s condition or symptoms* remained the highest ranked. *Early interventions options* ranked 10 when first diagnosed but decreased in ranking to 17 at the later time point.

Other topics increased in the rankings, for example, *managing family dynamics* ranked at 20 at the time of diagnosis compared with 16 at present. The greatest difference in ranking (7) was for *educational options*, which ranked 16 at time of diagnosis, but at present ranked 9. Similarly, *upcoming events or workshops* increased in ranking by 8 from 19 at the time of diagnosis to 11 at present. Interestingly *future pregnancies* changed little across timelines, ranking 15 at the time of diagnosis versus 17 at present**.**

Impact

Having considered respondents’ usage of the Internet, next the respondents were asked whether Internet-sourced information influenced decisions they made about their child’s condition, and 98 of them replied. Most respondents considered that

Internet-sourced information had some degree of influence, be it a *minor* or *some influence* (67%, 66/98) or a *major influence* (20%, 20/98), whereas for 10% (10/98), it had *no influence* and 2% (2/98) *did not know*.

Parents were asked about the impact on them of information found on the Internet by choosing relevant statements from a list, and they could choose as many as applied to them. In total, 86 parents replied to this question, and their choices are ranked in the order of popularity in Table 6. For 72% (62/86) of respondents, Internet-sourced information improved their understanding of their child's condition. The next 2 highly ranked options were *enabled me to explain my child's condition* (58%, 50/86) and *improved my ability to manage and care for my child's condition* (57%, 49/86). For approximately a third of respondents (33%, 28/86), the impact of Internet-sourced information had served to increase their anxiety, while for 16% (14/86), it decreased their anxiety.

Table 6. The impact of Internet-sourced information on parents (n=86; Participants could choose as many options as applied to them)

Rank	Impact	n (%)
1	Improved my understanding of my child's condition	62 (72)
2	Enabled me to explain my child's condition	50 (58)
3	Improved my ability to manage and care for my child's condition	49 (57)
4	Increased my anxiety	28 (33)
5	Was useful for diagnosing my child's condition	20 (23)
6	Decreased my anxiety Was useful for accessing medicines or alternative treatments or therapies on the Internet	14 (16)
7	Made me change my medical or health care professional	8 (9)
8	Made me change my child's food habits	6 (7)
9	Made me change my child's physical activity	5 (6)
10	Was not useful Not sure	4 (5)
11	Empowered me	2 (2)
12	Helped me educate my health care professional Useful for making new contacts	1 (1)

Parents Disseminating Internet-Sourced Information

Respondents used the Internet to communicate and link with others, for example, via email, Facebook, Twitter, and so forth (Table 4). More specifically, the respondents were asked whether they were registered in a Web-based forum or a social network group dedicated to their child's condition, and of the 112 who answered this question, 80.4% (90/112) were. Of these 90 respondents, 86% (77/90) shared information about their child's condition with these Internet communities.

In the context of health care, the respondents were asked whether they told their doctor or health care professional about the information they found on the Internet regarding their child's condition. Seventy-six (78%, 76/97) who answered this question had relayed Internet-sourced information about their child, but there was variability regarding the degree of interest the doctor or health care professional had in the information. Of the 76 respondents, 50% (38/76) felt that their doctor or health care professional was *somewhat interested* in what they had found, whereas 22% (17/76) felt they were *very interested*, 16% (12/76) felt they were *not too interested*, 9% (7/76) felt that their doctor or health care professional was *not at all interested*, and 3% (2/76) *did not know*.

Discussion

The 3 objectives of this study were to (1) ascertain parents general Internet usage patterns, (2) identify the nature of the information parents most frequently searched for, and (3) determine the impact Internet-sourced information had on parents of children with rare conditions. All objectives were met and are discussed in the following sections.

Internet Use

All respondents were frequent and competent users of information and communication technology; all of them had an email account, and most (94.6%, 106/112) had a Facebook account. Respondents mostly searched the Internet from home, late in the evening, and were practiced at finding the information they sought; these patterns of general Internet usage concurred with similar literature [14,15].

This study found that parents were discerning when searching the Internet by comparing the content they found with their own experiences and knowledge. They took many factors into account such as the relevancy and accuracy of the information, how trustworthy and up to date it was, and who was disseminating the information; these judicious parent characteristics have been reported by others [14,16-18].

The parents in this study were well educated, and 90% of their children had a specific diagnosis. Despite this, respondents still accessed numerous sites seeking information regarding their child's condition, prognosis, and management and to ask questions, a finding supported by Roche and Skinner [19] and others [8,16,20-22].

This study sought data on respondents' information needs at 2 stages: at the time of their child's diagnosis and at present. Although there was some variation across the 2 timelines significantly, the highest ranked item remained *my child's condition or symptoms*. Similarly, there was consistency regarding the top 5 ranked information needs at the 2 timelines (Table 5). Pelentsov et al [23] have confirmed that despite the diverse characteristics of rare conditions, there is consistency regarding the common unmet needs of parents of children with rare conditions.

Impact

The most significant impact of Internet-sourced information was the empowering effect it had on parents, particularly their improved understanding of their child's condition. Internet-sourced information enabled them to explain their child's condition and improved their ability to manage and care for their child. Some studies reported that Internet-sourced information resulted in parents changing their health care professional [19,24,25], which was also the case for 9% parents in this study (Table 6).

Parents Disseminating Internet-Sourced Information

In this study, parents not only engaged with and consumed Internet-sourced information for their own needs, but they were active disseminators of newly sourced information about their child's condition to their health care professionals and wider support network. Social media, Facebook in particular, provided instant access to other parents of children with rare conditions. In this study, the majority (86%) of those registered in a relevant Web-based forum or network shared information about their child's condition. Traditionally peer support has been in the form of mailing lists, newsletters, discussion fora, and chat rooms [26]. However, it was evident in this study that Web-based engagement and exchange of information with other parents provided support and created a sense of belonging, which in turn reduced the feeling of exclusion, as has been reported by others [3,7,10,17,19,20,27-29].

Respondents were likely to talk to their health care professional about information they uncovered on the Internet, although not all health care professionals were interested in their findings. Few considered that the information had helped educate their health care professional, which is reported by others [30,31]. These findings were somewhat surprising, as even in 2002 doctors were being advised to be prepared to ask parents about their information needs and to discuss Internet use with them [32]. This lack of health care professional-parent exchange to determine what information parents are currently seeking and uncovering might be a missed opportunity in health care consultations and possibly improved child outcomes whereby health care professionals might better determine what services or supports parents require.

Implications

Implications for health care practice include the impact Internet-sourced information is having on parent-health care professional dynamics. These include how information is shared, health care professionals' interaction with parents when parents share with them the information they have sourced, and the subsequent changing nature of parents' management of their child's condition.

In this study, parents of children with rare conditions identified a number of key factors that should be considered when developing an Irish website. These include ensuring the following:

- The content is relevant, accurate, trustworthy, and up to date.
- The topics most frequently searched for (Table 5) are addressed.
- It contains a Web-based forum or a social network component.
- The website is integrated with social media and is mobile friendly.

These are fundamental factors considering that Internet-sourced information directly influences and impacts on parents (Table 6) and parents disseminate information widely to their health care professionals and to other parents and relevant networks.

Limitations

As this was an Internet survey and participation was voluntary, it was not representative of all parents of children with rare conditions and most likely attracted participation from parents who were competent Internet users. The questionnaire consisted of 45 questions, and parents who had more than 1 child with a rare condition were asked to complete another 8 questions per additional child (Table 1). The length of the questionnaire might have contributed to the attrition of respondents as they progressed through the questionnaire. As the questionnaire was available on the Internet, it was accessed and completed by parents residing outside of Ireland (Table 2). Parents were asked which Internet-sourced information they sought during 2 phases in their child's life: at the time of diagnosis and at present. However, this question was posed at 1 timepoint (at present) and parents were asked to retrospectively recall what information they were seeking when their child was first diagnosed. Nonetheless, this study did provide an insight into what parents used the Internet for and its impact. Further research into the relevance of information available to parents of children with rare conditions, the impact of Internet-sourced information on parents, the health care professionals' response to parents, and the use of social media in parent-to-parent support is recommended.

Conclusions

The findings of this study add to the body of emerging research that gives an insight into the use of, and reason for using, the Internet and information and communications technology by parents of children with rare conditions and the ever evolving parent-to-parent support systems (social and intellectual) via social media. What appears evident is that Internet and social media engagement facilitates the emergence of parents who are better informed and empowered, have greater understanding of the management and care of rare conditions, and are increasingly considered experts in their child's care, specifically in how the particular condition is developing in their child [6]; many parents come prepared to health consultations with information sourced from the Internet [1].

Acknowledgments

This study was funded by the Saoirse Foundation. The research team is most appreciative of the parents who participated in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

[PDF File (Adobe PDF File), 73KB - jmir_v19i2e51_app1.pdf]

Multimedia Appendix 2

[PDF File (Adobe PDF File), 36KB - jmir_v19i2e51_app2.pdf]

References

1. Fergie G, Hilton S, Hunt K. Young adults' experiences of seeking online information about diabetes and mental health in the age of social media. Health Expect 2015 Dec 08;19(6):1324-1335. [doi: 10.1111/hex.12430] [Medline: 26647109]
2. van den Bree MB, Miller G, Mansell E, Thapar A, Flinter F, Owen MJ. The internet is parents' main source of information about psychiatric manifestations of 22q11.2 deletion syndrome (22q11.2DS). Eur J Med Genet 2013 Aug;56(8):439-441 [FREE Full text] [doi: 10.1016/j.ejmg.2013.05.001] [Medline: 23707654]
3. Pelentsov LJ, Laws TA, Esterman AJ. The supportive care needs of parents caring for a child with a rare disease: a scoping review. Disabil Health J 2015 Oct;8(4):475-491. [doi: 10.1016/j.dhjo.2015.03.009] [Medline: 25959710]
4. Dragusin R, Petcu P, Lioma C, Larsen B, Jørgensen HL, Cox IJ, et al. FindZebra: a search engine for rare diseases. Int J Med Inform 2013 Jun;82(6):528-538. [doi: 10.1016/j.ijmedinf.2013.01.005] [Medline: 23462700]
5. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
6. Nicholl H, Doyle C, Begley T, Murphy M, Lawlor A, Malone H. Developing an information leaflet on 22q11.2 deletion syndrome for parents to use with professionals during healthcare encounters. J Spec Pediatr Nurs 2014 Jul;19(3):238-246. [doi: 10.1111/jspn.12078] [Medline: 24674229]
7. Anderson RP, Elliott EJ, Zurynski YA. Australian families living with rare disease: experiences of diagnosis, health services use and needs for psychosocial support. Orphanet J Rare Dis 2013 Feb 11;8:22 [FREE Full text] [doi: 10.1186/1750-1172-8-22] [Medline: 23398775]
8. Nieuwboer CC, Fukkink RG, Hermanns JM. Peer and professional parenting support on the Internet: a systematic review. Cyberpsychol Behav Soc Netw 2013 Jul;16(7):518-528. [doi: 10.1089/cyber.2012.0547] [Medline: 23659725]
9. Binford Hopf RB, Le Grange D, Moessner M, Bauer S. Internet-based chat support groups for parents in family-based treatment for adolescent eating disorders: a pilot study. Eur Eat Disord Rev 2013 May;21(3):215-223. [doi: 10.1002/erv.2196] [Medline: 22949134]
10. Paterson BL, Brewer J, Stamler LL. Engagement of parents in on-line social support interventions. J Pediatr Nurs 2013 Apr;28(2):114-124. [doi: 10.1016/j.pedn.2012.05.001] [Medline: 22710151]
11. Porter A, Edirippulige S. Parents of deaf children seeking hearing loss-related information on the internet: the Australian experience. J Deaf Stud Deaf Educ 2007;12(4):518-529 [FREE Full text] [doi: 10.1093/deafed/enm009] [Medline: 17456507]
12. Tozzi AE, Mingarelli R, Agricola E, Gonfiantini M, Pandolfi E, Carloni E, et al. The internet user profile of Italian families of patients with rare diseases: a web survey. Orphanet J Rare Dis 2013;8:76 [FREE Full text] [doi: 10.1186/1750-1172-8-76] [Medline: 23680013]
13. Massey OT. A proposed model for the analysis and interpretation of focus groups in evaluation research. Eval Program Plann 2011 Feb;34(1):21-28. [doi: 10.1016/j.evalprogplan.2010.06.003] [Medline: 20655593]
14. Hand F, McDowell DT, Mc Dowell DT, Glynn RW, Rowley H, Mortell A. Patterns of internet use by parents of children attending a pediatric surgical service. Pediatr Surg Int 2013 Jul;29(7):729-733. [doi: 10.1007/s00383-013-3317-5] [Medline: 23615872]

15. Rudi J, Dworkin J, Walker S, Doty J. Parents' use of information and communications technologies for family communication: differences by age of children. Inf Commun Soc 2014 Jul 02;18(1):78-93. [doi: 10.1080/1369118X.2014.934390]
16. Oprescu F, Campo S, Lowe J, Andsager J, Morcuende JA. Online information exchanges for parents of children with a rare health condition: key findings from an online support community. J Med Internet Res 2013;15(1):e16 [FREE Full text] [doi: 10.2196/jmir.2423] [Medline: 23470259]
17. Glenn AD. Using online health communication to manage chronic sorrow: mothers of children with rare diseases speak. J Pediatr Nurs 2015;30(1):17-24 [FREE Full text] [doi: 10.1016/j.pedn.2014.09.013] [Medline: 25316198]
18. Ladd D. General and health information challenges of patients with rare diseases: the importance of health information provision and web sites for locating rare disease resources. J Hosp Librariansh 2015 May 05;15(2):189-197. [doi: 10.1080/15323269.2015.1014762]
19. Roche MI, Skinner D. How parents search, interpret, and evaluate genetic information obtained from the internet. J Genet Couns 2009 Apr;18(2):119-129 [FREE Full text] [doi: 10.1007/s10897-008-9198-4] [Medline: 18937062]
20. Kingsnorth S, Gall C, Beayni S, Rigby P. Parents as transition experts? Qualitative findings from a pilot parent-led peer support group. Child Care Health Dev 2011;37(6):833-840. [doi: 10.1111/j.1365-2214.2011.01294.x] [Medline: 22007983]
21. Niela-Vilén H, Axelin A, Salanterä S, Melender HL. Internet-based peer support for parents: a systematic integrative review. Int J Nurs Stud 2014 Nov;51(11):1524-1537. [doi: 10.1016/j.ijnurstu.2014.06.009] [Medline: 24998788]
22. Nölke L, Mensing M, Krämer A, Hornberg C. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. BMC Public Health 2015 Jan 29;15(1):31 [FREE Full text] [doi: 10.1186/s12889-015-1423-0] [Medline: 25631456]
23. Pelentsov LJ, Fielder AL, Laws TA, Esterman AJ. The supportive care needs of parents with a child with a rare disease: results of an online survey. BMC Fam Pract 2016 Jul 21;17(1):88 [FREE Full text] [doi: 10.1186/s12875-016-0488-x] [Medline: 27439905]
24. Leonard H, Slack-Smith L, Phillips T, Richardson S, D'Orsogna L, Mulroy S. How can the Internet help parents of children with rare neurologic disorders? J Child Neurol 2004 Nov;19(11):902-907. [doi: 10.1177/08830738040190110901] [Medline: 15658796]
25. Skinner D, Schaffer R. Families and genetic diagnosis in the genomic and Internet age. Infants Young Child 2006;19(1):16-24 [FREE Full text]
26. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [Medline: 12020305]
27. McGarvey B, Hart C. Rehab. 2008. An investigation into the social support needs of families who experience rare disorders on the island of Ireland URL: http://www.rehab.ie/about/PDFS/July2008/RehabCare_RD_Report.pdf [accessed 2017-02-15] [WebCite Cache ID 6oJ2cW1Eu]
28. Gundersen T. 'One wants to know what a chromosome is': the internet as a coping resource when adjusting to life parenting a child with a rare genetic disorder. Sociol Health Illn 2011 Jan;33(1):81-95 [FREE Full text] [doi: 10.1111/j.1467-9566.2010.01277.x] [Medline: 20937053]
29. Hall JG. The role of patient advocacy/parent support groups. S Afr Med J 2013 Oct 11;103(12):1020. [doi: 10.7196/samj.6976]
30. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
31. Schieppati A, Henter JI, Daina E, Aperia A. Why rare diseases are an important medical and social issue. The Lancet 2008 Jun 14;371(9629):2039-2041. [doi: 10.1016/S0140-6736(08)60872-7] [Medline: 18555915]
32. Tuffrey C, Finlay F. Use of the internet by parents of paediatric outpatients. Arch Dis Child 2002 Dec;87(6):534-536 [FREE Full text] [Medline: 12456558]

Edited by G Eysenbach; submitted 05.04.16; peer-reviewed by D McDowell, E Agricola; comments to author 27.07.16; revised version received 28.10.16; accepted 13.12.16; published 28.02.17.

Please cite as:
Nicholl H, Tracey C, Begley T, King C, Lynch AM
Internet Use by Parents of Children With Rare Conditions: Findings From a Study on Parents' Web Information Needs
J Med Internet Res 2017;19(2):e51
URL: http://www.jmir.org/2017/2/e51/
doi:10.2196/jmir.5834
PMID:28246072

Original Paper

Searching for Real-World Effectiveness of Health Care Innovations: Scoping Study of Social Prescribing for Diabetes

Karen Pilkington[1,2], BPharm (Hons), MSc, PhD; Martin Loef[3], PhD; Marie Polley[1], BSc (Hons), PhD

[1]School of Life Sciences, Faculty of Science and Technology, University of Westminster, London, United Kingdom

[2]School of Health Sciences and Social Work, Faculty of Science, University of Portsmouth, Portsmouth, United Kingdom

[3]Institute of Transcultural Health Studies, European University Viadrina, Frankfurt, Germany

Corresponding Author:
Karen Pilkington, BPharm (Hons), MSc, PhD
School of Health Sciences and Social Work
Faculty of Science
University of Portsmouth
St Michael's Building
2 King Richard 1st Road
Portsmouth, P01 2FR
United Kingdom
Phone: 44 123 984 4497
Email: karen.pilkington@port.ac.uk

Abstract

Background: Social prescribing is a process whereby primary care patients are linked or referred to nonmedical sources of support in the community and voluntary sector. It is a concept that has arisen in practice and implemented widely in the United Kingdom and has been evaluated by various organizations.

Objective: The aim of our study was to characterize, collate, and analyze the evidence from evaluation of social prescribing for type 2 diabetes in the United Kingdom and Ireland, comparing information available on publicly available websites with the published literature.

Methods: We used a broad, pragmatic definition of social prescribing and conducted Web-based searches for websites of organizations providing potentially relevant services. We also explored linked information. In parallel, we searched Medline, PubMed, Cochrane Library, Google Scholar, and reference lists for relevant studies published in peer-reviewed journals. We extracted the data systematically on the characteristics, any reported evaluation, outcomes measured and results, and terminology used to describe each service.

Results: We identified 40 UK- or Ireland-based projects that referred people with type 2 diabetes and prediabetes to nonmedical interventions or services provided in the community. We located evaluations of 24 projects; 11 as published papers, 12 as Web-based reports, and 1 as both a paper and a Web-based report. The interventions and services identified included structured group educational programs, exercise referral schemes, and individualized advice and support with signposting of health-related activities in the community. Although specific interventions such as community-based group educational programs and exercise referral have been evaluated in randomized controlled trials, evaluation of individualized social prescribing services involving people with type 2 diabetes has, in most cases, used pre-post and mixed methods approaches. These evaluations report generic improvement in a broad range of outcomes and provide an insight into the criteria for the success of social prescribing services.

Conclusions: Our study revealed the varied models of social prescribing and nonmedical, community-based services available to people with type 2 diabetes and the extent of evaluation of these, which would not have been achieved by searching databases alone. The findings of this scoping study do not prove that social prescribing is an effective measure for people with type 2 diabetes in the United Kingdom, but can be used to inform future evaluation and contribute to the development of the evidence base for social prescribing. Accessing Web-based information provides a potential method for investigating how specific innovative health concepts, such as social prescribing, have been translated, implemented, and evaluated in practice. Several challenges were encountered including defining the concept, focusing on process plus intervention, and searching diverse, evolving Web-based sources. Further exploration of this approach will inform future research on the application of innovative health care concepts into practice.

(J Med Internet Res 2017;19(2):e20) doi:10.2196/jmir.6431

KEYWORDS

diabetes mellitus, type 2; evaluation studies; primary health care; program evaluation

Introduction

The World Health Organization global report on diabetes revealed that the number of people with diabetes had risen from 108 million in 1980 to 422 million in 2014 [1]. A similarly dramatic increase is anticipated in the future in the United Kingdom, as the number of people with type 2 diabetes is projected to increase by 50% between 2010 and 2030 [2]. Major advances have occurred in diabetes research, showing that lifestyle interventions can delay or prevent the onset of diabetes [3,4], but prevention and good control of diabetes remain elusive for most of the population [5]. The primary care sector represents a potentially important setting in this context, as more than 85% of the population in the United Kingdom visits a general practitioner (GP) at least once a year [6]. Furthermore, to increase the sustainability of general practice, other options for the support and care of patients, particularly those with chronic conditions, are being sought by organizations such as National Health Service (NHS) England [7].

Social prescribing (or “community referral”) is a relatively new approach in health care, aiming to create referral pathways that enable the GP or a health care practitioner to refer patients with social or practical needs to a local provider of nonclinical services [8,9]. These are often offered by volunteers or the community sector and cover a wide range of interventions including educational sessions, exercise training, dietary advice, creative activities, self-help groups, emotional or social support, and stress management. A current challenge is to systematically collate and evaluate the evidence of the impact of social prescribing on people’s lives, its conceptualization, dissemination, and the way it is operationalized in practice.

In general, the translation of research into practice encompasses multiple dimensions, with different levels of scientific involvement and routes of publication. It covers projects that analyze selected aspects of the implementation of new treatments or research knowledge into a practical setting to ensure they not only reach the patients for whom they are intended but are also implemented in the most effective manner [10]. The focus in research is frequently on the efficacy and generalizability of public health interventions, whereas there may be equal levels of interest in the local implementation of interventions in a real-world setting, with “scientific” outcomes being of minor interest, resulting in a low priority for publication. As a result, using the standards of and processes for conducting systematic reviews on health care interventions [11,12] may not represent an adequate approach for gathering evidence of translational research including real-world projects. Limited numbers of studies are available in peer-reviewed articles and, therefore, in the major databases for synthesizing data and drawing conclusions. However, although other sources such as evaluations in practice may be available, these may be limited in methodological and reporting quality and difficult to locate.

Previous reviews of the evidence on social prescribing have been reported [13,14] as have evaluations of schemes such as exercise on referral [15-17], which fit the concept of social prescribing. Innovative care programs have been evaluated, which include elements of social prescribing and encourage commissioning of nontraditional providers to support people with long-term conditions including diabetes [18], as have fitness programs involving various initiatives across the United Kingdom [19]. There has not, however, been an in-depth review focusing on social prescribing for a specific health area such as diabetes in a specific context (the United Kingdom) with a comparison of the data available from formal (published) and informal Web-based (gray literature) sources.

The aim of this study was to characterize social prescribing, and to then collate and analyze evaluations of social prescribing, by focusing on services or projects for people with type 2 diabetes in the United Kingdom and Ireland. This allowed an examination of how an innovative health care concept has been translated and applied in practice and provided an indication of the evidence that is available. By searching publicly available websites in addition to the published literature, it also allowed a comparison of the information available from formal (published literature) and informal (websites and related Web-based information) sources, and an exploration of the potential value, feasibility, and challenges presented by such an approach to searching for evidence.

Methods

Overview

The overall approach is based on that of a scoping review or study, a framework for which was provided by Arksey and O’Malley [20]. The following elements of the framework, in particular, informed the approach: identifying all literature regardless of study design, redefining search terms as familiarity with the literature increases, and using an iterative process. However, our methods were also guided by our aim to explore information available on websites about real-world projects or services and compare this with the published literature.

Search Strategies

Because of the broad scope and varied terminology associated with social prescribing, and the intention to explore the gray literature in additional to conventional databases, we used a range of search strategies and an iterative approach. We searched the Cochrane Library, MEDLINE, PubMed, and Google scholar, and various sources of gray literature (Google, Yahoo, Bing, greylit.org, Opengrey.eu, and specific NHS websites such as NHS Evidence).

The initial search for published studies combined the following thematic fields to identify social prescribing projects: (1) social prescri* OR refer*; (2) program OR treatment OR management OR education OR support OR physical exercise OR aerobic

OR physical activity OR leisure-time OR exercise OR sport OR leisure activit* OR physical fitness OR training OR physical performance OR weight loss OR weight reduction OR BMI OR body weight OR body mass index OR obesity OR overweight OR adiposity OR smoking OR tobacco OR cigarette OR social support OR loneliness; (3) diabetes; (4) primary care OR GP OR general practitioner OR community OR voluntary. In addition, we performed a search on lifestyle intervention trials for type 2 diabetes in MEDLINE and searched the 200 most recent hits for studies on social prescribing to test the effectiveness of our search strategy.

The search for gray literature (websites and related Web-based material) was based on the combination and iterative modification of the terms “social prescribing,” “social prescription,” ”primary care,” ”diabetes,” ”referral,” and “community services.” For instance, we searched the first 100 hits in the search engines for “social prescribing” and “diabetes.” If we found NHS websites with links or references to other projects of potential relevance, we followed these; otherwise we modified the search terms and repeated the process multiple times. An initial set of searches was conducted in December 2014.

The aforementioned searches were repeated in October 2016. At this point, further searches were conducted. Three search strategies for databases were developed based on the results of previous searches including further terms for link workers and “social prescribing interventions” and a term for prediabetes based on the rationale that the interventions would be of a similar nature to those for type 2 diabetes. These strategies are presented in Multimedia Appendix 1. The final search was a further Web-based search via the Google.co.uk search engine using the strategy “social prescribing” AND diabetes AND evaluation. The inclusion of evaluation as a term in this search was due to the fact that in the interim since the previous search, the use of the term social prescribing on websites had become much more extensive. Therefore, it was necessary to focus on identifying projects that had been evaluated. Searches for each of the individual projects located on websites were also carried out by searching using the project name on PubMed.

Selection Criteria

Various definitions of social prescribing have been proposed, but it is clear that there is still a lack of consensus on its scope and interpretation. Therefore, a broad, pragmatic set of criteria was used to identify studies or projects that were likely to fit within the spectrum of relevant interventions: located in the United Kingdom or Ireland and involving a primary care provider referring patients including those with type 2 diabetes or prediabetes to a third party that is delivering nonmedical services in the community. Studies or evaluations had to include a description of the referral process or be described as a social prescribing service and mention type 2 diabetes or prediabetes either in the inclusion criteria or in the report or publication. In line with the aim of a scoping review, all study designs used in evaluation were included as were all forms of report. There was no formal language restriction, but studies and projects were excluded if they were available only as abstracts, protocols, located outside the United Kingdom or Ireland, or did not provide details regarding the process of referring the patients from primary care to the external health care, community, or voluntary service. The searches were conducted by 2 authors (ML and KP). Initial screening of each set of search results was carried out by 1 author, preliminary selections by each of the 2 authors were then compared to compile a “short-list,” full-texts were consulted, and final selections were made with reference to the third author (MP) where necessary.

Data Extraction and Analysis

For each service, data were extracted by 1 of the authors (ML or KP) into a Microsoft Excel spreadsheet, documenting the name of the service, the name of the organization that carried out the service, the type of organization, the location, the criteria for referral, the type of intervention, outcomes, or objectives, the evaluation of the projects and its findings, the type of publications, and the terminology used in the project regarding social prescription. All extracted data was checked by at least one other author (KP, ML, or MP). For those projects that had undergone an evaluation where the relevant published papers or Web-based report was located, details of the evaluation were extracted. As a range of study designs were used in evaluation and because the main aim was to scope the information available, full appraisal was not carried out, as is the usual approach at this stage [21]. To provide an indication of the quality of the evidence on effectiveness, the following items were extracted: study design (including whether randomized and, where relevant, the control interventions employed), sample size, data collection methods, and outcomes measured. For this purpose, the main study or latest evaluation report was accessed. Details of any additional or related studies are discussed in the text.

Results

Search Results and Study Characteristics

A total of 40 projects or services and 24 evaluations were identified through the various searches and sources. A total of 3249 records from databases were screened, and a “shortlist” of 302 potentially relevant studies was compiled. After further screening and exclusion of studies, 29 papers related to 12 projects were selected, which met the inclusion criteria. For the Web-based searches, in addition to iterative searching, 189 documents retrieved by the final Google search were accessed, resulting in a list of 40 possibly relevant projects, of which 11 were excluded (2, not diabetes; 3, only type 1 diabetes; 3, no clear referral from primary care; 3, complex change interventions including but not focused on social prescribing). Twenty-nine projects met the inclusion criteria, with evaluations found for 13 projects (a Web-based report and a published paper were found for 1 project). Three additional project evaluations were located that included people with long-term conditions but, as diabetes was not specifically mentioned, these were not included. The screening process is shown in Figure 1.

Summaries of the evaluated projects or services focusing on the key study or report in each case are shown in Table 1. Details of the additional projects located on the Internet are presented in Table 2. The references to papers, reports, and websites are included for each project or service.

Figure 1. Process for the identification, screening, and selection of projects and evaluations.

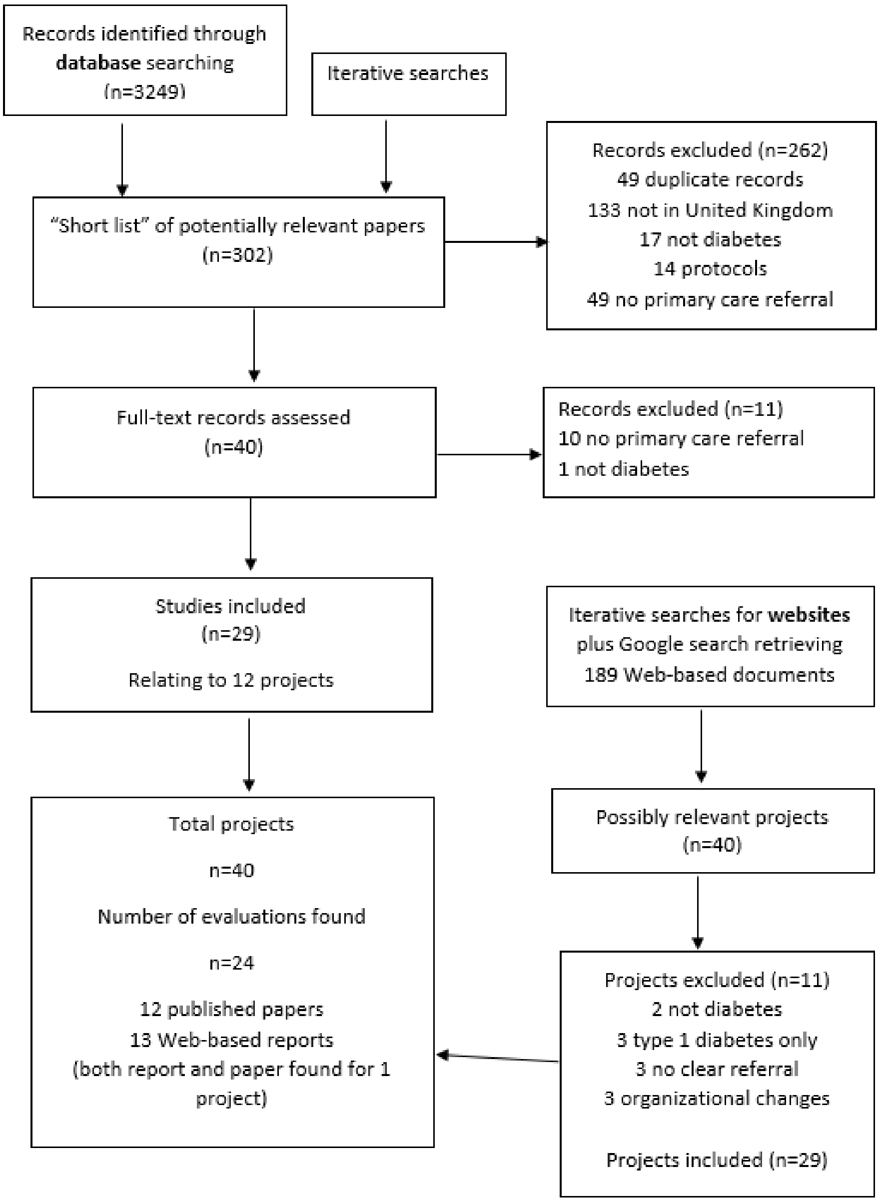


Table 1. Evaluated projects or services involving nonmedical community-based interventions.

Project or service	Intervention	Design of study or evaluation	Participants and setting	Outcomes and findings	Source
Addition-Plus [22,23]	Individually tailored, behavioral change intervention delivered by lifestyle facilitators plus intensive treatment	RCT[a] (compared with intensive treatment alone) Also 5-year follow-up	478 adults with T2DM[b] referred by HCP[c] at 34 general practices in Eastern England (239 per group)	Health behaviors and cardiovascular risk factors (physical activity, diet, and smoking status) improved but no significant difference between groups after 1 year (*P*=.36, *P*=.07, and *P*=.28, respectively)	PP[d]
Age UK's Fit for the Future "Social Prescribing" extension project [24,25]	Social prescribing project (health care professional referral and range of support)	Service evaluation (mixed methods for main project; surveys and qualitative interviews for this extension project)	305 people with LTCs[e] from 3 areas across the United Kingdom completed pre-post questionnaires; (247 completed both time points) (diabetes 16% of baseline group)	At 3-month follow up: less significant improvements in mental well-being (*P*=.36), attitude to healthy eating (*P*=.16) and social networks compared with the main project that varied across the 3 areas but some positive changes for life satisfaction; attitude to physical activity; and exercising	WR[f]
Birmingham Exercise on Prescription [26-28]	Exercise referral intervention grounded in Self-Determination Theory (SDT)	Exploratory cluster RCT (compared with standard exercise referral alone) Also regression analysis of support	347 adults with risk factors for CHD[g] including T2DM referred to an exercise scheme at 13 leisure centers in Birmingham (184 in SDT group; 163 in standard group); number with T2DM not reported	Primary outcome: Physical activity using the 7-Day Physical Activity Recall (7DPAR). Other outcomes: BP[h], BMI[i], general health and fitness, anxiety, depression, vitality, quality of life, and well-being Improvements in physical activity, quality of life, and well-being in both groups. Between-group changes only significant for anxiety at 6 months (*P*=.003); physical activity (*P*=.93); quality of life (*P*=.40)	PP
City and Hackney Social Prescribing Project [29-32]	Social prescribing project (referral by GP[j] to social prescribing coordinators managed by a voluntary, community, and social enterprise service for personalized signposting)	Mixed methods evaluation with control group (23 GP practices in total; 6 control GP practices) (survey or in-depth interviews)	184 people with depression, anxiety, or T2DM supported by project based in a London borough (matched with 302 in control practices). Number with T2DM not reported	Reported no statistically significant change in health, well-being, anxiety, depression, or A + E visits due to the SP intervention at 8 months. Qualitative interviews revealed positive or extremely positive experiences	WR
Community Oriented Diabetes Education (CODE or CODET2) [33-35]	Structured self-management education program	Mixed methods pre-post evaluation (physiological tests, questionnaires, and semi-structured interviews)	401 adults with T2DM referred to 31 CODE programs across the Republic of Ireland; 392 completed baseline; 237 (60%) completed the post-program evaluation	No difference in HbA1c[k] (*P*=.14) or cholesterol (*P*=.06). Significant positive change in weight loss (*P*<.001), Significant increase in empowerment (*P*=.047), knowledge (*P*=.01) and quality of life (*P*<.001) at 26 weeks	PP and WR
DESMOND Newly Diagnosed [36-47]	6-hour structured group education program delivered by trained educators	Multicenter cluster RCT (compared with usual care) Also qualitative studies, 3 year follow-up, and a cost-effectiveness analysis	824 adults with T2DM referred to trial from 207 general practices in 13 primary care sites in United Kingdom	No significant difference in HbA1c (*P*=.52). Significant difference in weight loss (*P*=.03), smoking cessation (*P*=.03), changes in illness belief scores (*P*=.001), and depression (*P*=.03) after 12 months post diagnosis, in intervention group	PP
DESMOND Let's Prevent Diabetes [48,49]	6-hour group structured education program	Cluster RCT (compared with standard care) Also retrospective analysis	880 adults with prediabetes from 44 GP practices in Leicestershire (invited by GP)	Nonsignificant 26% reduced risk of developing T2DM in the intervention arm (*P*=.18). Significant improvements in HbA1c (*P*<.05), LDL cholesterol (*P*<.05), sedentary time (*P*<.01), and step count (*P*<.01) at 3 years	PP

Project or service	Intervention	Design of study or evaluation	Participants and setting	Outcomes and findings	Source
DESMOND Walking Away from Diabetes [47,50]	3-h group-based structured education program with pedometer use	Cluster RCT (compared with information provision)	808 people at high risk of T2DM from 10 GP practices in Leicestershire (data on 571 (71%))	Increases in ambulatory activity (*P*=.006) and self-reported vigorous-intensity physical activity at 12 months. No differences between groups at 3 years (*P*=.52) or in cardiometabolic markers	PP
EDIPS (European Diabetes Prevention Study)-Newcastle [51-53]	Intensive behavioral interventions to promote dietary modification and increased physical activity (group sessions plus signposting to community physical activities)	RCT (compared with usual care) Also qualitative study and follow-up	102 people with impaired glucose tolerance (prediabetes) (51 per group) referred by GP to trial and then to community physical activities in Newcastle	Significant reduction in the risk of developing T2DM, RR^m 0.45 (95% CI 0.2-1.2). Interim benefit achieved but no significant differences in sustained change (more than 2 years) for other outcomes (change in physical activity, fat, fiber, and carbohydrate intake). Mean duration of follow-up was 3.1 years	PP
EXERT (Exercise Evaluation Randomised Trial) [54]	Exercise referral scheme	3-arm RCT (compared leisure center-based exercise program, an instructor-led walking program, and advice-only)	943 patients from GP practices in 1 London borough; 13% diabetes (number with T2DM not reported)	Changes in exercise behavior and cardiovascular risk factors; waist–hip ratio, BMI, body fat, fitness, lifestyle behaviors, health status, quality of life, and health service usage and cost. All groups improved with no consistent differences between groups	PP
Living Well, Taking Control program [55,56]	Community-based program for diabetes prevention and newly diagnosed T2DM	Service evaluation involving economic evaluation, pre-post evaluation with all participants, process evaluation plus RCT (ComPoD) compared against wait list control group)	223 participants with T2DM (32.8%), 448 with prediabetes (66.0%) in service evaluation, and 40 recruited to trial in Birmingham and Bristol at the point of preliminary report	Preliminary results only: 6-month follow-up data for 123 participants. Significant pre-post changes in diabetes risk factors: weight loss (*P*<.001); HbA1c (*P*<.001); fiber intake (*P*<.005), depressive symptoms (*P*<.001); general health state (*P*=.002)	WR
Newly Diagnosed Type 2 Diabetes Dietary Education Group [57,58]	Group education program	Pre-post questionnaire to assess knowledge	126 adults with T2DM referred to program by GPs in North Wales	Significant differences in percentage of correct answers pre-post (*P* values <.01 for 9 of 11 questions)	WR
Newcastle Social Prescribing Project or NESTA People Powered Health project [59]	Social prescribing project (referral by primary care staff to nonclinical community services and networks plus information resource)	Mixed methods service evaluation (consultation with key stakeholders; assessment of project information and data; interviews with the health care professionals and patients)	124 people with LTCs referred from 6 organizations. Diabetes mentioned in cases only	Evaluation was based on numbers of patients achieving their individual goal (mainly health-related) and views of patients and HCPs on the service. Findings were used to inform a larger-scale project (ongoing).	WR
New Life New You [60-63]	Community-based lifestyle intervention; self-referral and signposting from primary care	Mixed methods pilot study (uncontrolled before-and-after study design with embedded interviews). Also interviews with black and minority ethnic community	218 people with impaired glucose tolerance (prediabetes) referred to project in Middlesbrough	Beneficial changes in physical activity, weight and waist measurements, and Finnish Diabetes Risk Score (FINDRISC) at 12 months. Follow-up with 134 (61%) participants	PP

Project or service	Intervention	Design of study or evaluation	Participants and setting	Outcomes and findings	Source
PoLLeN (People, Life, Landscape and Nature) Bromley By Bow [64]	Social and therapeutic horticulture project	Service evaluation (mixed methods approach with validated outcomes questionnaires, feedback, interviews, and case studies)	39 adults with mental distress and conditions including diabetes referred to program in a London borough. One case study mentions diabetes	Short form of Clinical Outcomes in Routine Evaluation questionnaire (CORE10) and SWEMWBS[l]. Varied numbers of participants completed questionnaires at 3 time points. No significant changes pre-post. Qualitative data from clients showed appreciation for the project	WR
Ramadan Education and Awareness in Diabetes (READ) [65]	Ramadan-focused education program delivered by ethnic-speaking HCP and community link worker	Retrospective analysis of 2 groups (A and B)	111 Muslim adults with T2DM in Brent, London (Group A: 57 people referred or self-referred to and attended program. Group B: 54 invited, did not attend)	Significant differences in weight loss (*P*<.001), and hypoglycemic events (*P*=.001), which were sustained in group A at 12 months	PP
Rotherham Social Prescribing service [66-68]	Social prescribing project (individual advice, signposting service from Voluntary and Community services (VCS) advisors)	Service evaluation (monitoring data, interviews, case studies, and surveys)	1,607 people with LTCs referred to the service. Diabetes only referred to in case studies	Reduction in demand for hospital care, improvements in well-being, and social impact were reported based on estimates from a subset of beneficiaries and cost-saving estimated	WR
Rugby Social Prescribing Project: ConnectWELL [69]	Pilot social prescribing project in 4 GP surgeries to support and signpost individuals to services and activities in the local and voluntary community	Service evaluation aimed to produce statistical evidence and recommendations for other social prescribing initiatives	People in Rugby with various health problems, for example, diabetes. Number with T2DM not mentioned	Interim report located online but publication status unclear. Anticipate data reporting in full report (as yet not found). No T2DM specific-outcomes reported	WR
Sadee Smile (South Asian Diabetes Education, Empowerment and Self-Management in Leeds) [70]	Pilot education program led by nonclinical tutors	Service evaluation of 11 courses (pre-post) including knowledge questionnaire, focus groups, case studies, and interviews with staff	113 adults with T2DM referred to program in Leeds	Knowledge, skills, and confidence assessed around diabetes management improved post service. Increased physical activity reported	WR
South Gloucestershire Exercise on Prescription [71-72]	Tailored, supervised exercise referral scheme. Patients with LTCs referred by GP	Pre-post comparison of physiological data, WEMWBS, demographics on participants, and usage data; cost data; interviews with staff and patients	2505 participants referred to scheme. Number with T2DM: 21.9% of males and 9.9% of females.1379 (55%) considered completers	At 12-week follow-up: Significant decrease in systolic BP (*P*<.01), waist (*P*<.03), number of 30-min physical activity sessions and well-being (*P*<.001). No significant difference in weight (*P*=.63), BMI (*P*=.25), hip (*P*=.86), or diastolic BP (*P*=.56)	WR
Well UK South West Well-being or South West Well-being (SWWB) [73,74]	Social prescribing project (including a portfolio of initiatives)	Service evaluation (longitudinal study using outcome measures, surveys and interviews, case studies)	737 people with low-level mental ill health, approaching older age, and families on lower incomes referred from Bristol area. Case studies mention diabetes	Positive changes in general health, physical activity, diet, mental well-being, and social well-being based on self-reporting	WR

Project or service	Intervention	Design of study or evaluation	Participants and setting	Outcomes and findings	Source
Wigan and Bolton Exercise Referral Scheme [75]	Exercise referral scheme plus information	RCT (compared with information alone)	545 sedentary adults with risk factors for coronary heart disease including diabetes from 46 of 52 general practices in one borough in northern England (275 intervention, 270 control). Number with T2DM not reported	Primary outcome: moderate or vigorous activity for 90 min per week: significant improvement at 6 months (*P*=.05) but not significant at 12 months (*P*=.18) based on 7-Day Physical Activity Recall (7dPAR)	PP
Wirral Lifestyle and Weight Management program [76]	Intensive 12-week program consisting of a variety of group meetings or tailored one-to-one sessions	Economic analysis (preliminary report)	3810 people with obesity with comorbidities of T2DM or CHD. Number with T2DM not reported	Majority of subgroups showed significant reductions in weight at 12 weeks (*P*<.001). Limited data available for all other outcomes except BMI. Associated cost savings for T2DM calculated	WR
X-PERT [77-79]	Patient-centered, group-based self-management program (6 2-hour sessions)	RCT (compared with individual appointments)	314 adults with T2DM in 3 boroughs in northern England (157 per group)	Significant improvement in X-PERT group compared with control for HbA1c, weight, BMI, waist circumference (all *P*<.001), total cholesterol (*P*<.01), self-empowerment (*P*<.04), knowledge (*P*<.001); also in physical activity, foot care, fruit and vegetable intake, enjoyment of food at 14 months	PP

[a] RCT: Randomized controlled trial.

[b] T2DM: type 2 diabetes.

[c] HCP: Health care practitioners.

[d] PP: Published journal paper.

[e] LTC: Long-term conditions.

[f] WR: Web-based report.

[g] CHD: Coronary heart disease.

[h] BP: Blood pressure.

[i] BMI: Body mass index.

[j] GP: General practitioner.

[k] HbA1c: Hemoglobin A1c.

[l] (S)WEMWBS: (Short) Warwick Edinburgh Mental Wellbeing Scale.

[m] RR: Risk ratio.

Table 2. Additional projects or services identified from Web-based searches.

Name of project or service	Location	Referral process	Types of patients	Type of service, what is offered and by whom	Aim of service
Altogether Better Diabetes[a][80,81]	Northern England (Leeds, York)	Community Health Champions (CHCs) based in GP[b] practices can refer to GP or to community	Diabetes	CHCs signpost clients to activities, accompany clients, and provide networking opportunities	Improving the health and well-being of communities
Be active plus [82,83]	Birmingham (South and Central)	GP or nurse refers patients at surgery to program	Different conditions including diabetes, hypertension, and obesity	Individually tailored exercise program plus support from Health and Fitness Advisors at local leisure center	To increase the amount of physical exercise
Building Health Partnerships: Bristol [84]	Bristol	Referral to Public Health Improvement Teams	Black and minority ethnic diabetic (primarily Somali and Asian people)	Individual support from Public Health Improvement Teams, cooking events	Healthier lifestyle
Camden Exercise Referral[a] [85-87]	London (Camden)	GP or health professional refers to Active Health team	A range of conditions including diabetes	Individualized exercise program	Active lifestyle changes
Diabetes Education and Revision in East Kent (DEREK) [88]	East Kent	GP or practice nurse refers patients to education coordinator	Type 2 diabetes	Education program at various venues	Better management of diabetes, shared experiences, better relationships with health care professionals
Diabetes Education Awareness for Life (DEAL)[a] [89]	Berkshire	GP or practice nurse or community nurse refers patient to program	Newly diagnosed and existing type 2 diabetes	Education program (also well-being groups for talking therapy) in community	Healthier lifestyle, education
Go4life [90-92]	North Somerset	Referral to volunteers that offer 1:1 support for up to 6 weeks	Type 2 diabetes	Individual support and education	Healthier lifestyle, specifically meeting physical activity goals
Good2go [93]	York	GP or practice nurse refers patient to program	Type 2 diabetes	Education program	To better manage the disease
HARRIET (Harrogate Initiative for education in type 2 diabetes) [94,95]	North Yorkshire and York	GP referral	Type 2 diabetes	Education program	Not specifically stated (general aim: to help people become experts on managing their condition')
Juggle Diabetes Education Service [96]	Nottingham	GP refers patient to program delivered in variety of venues (or can self-refer)	Type 2 diabetes (not treated with insulin)	Education program	To better manage the disease
King's Lynn diabetes type 2 education program [97]	King's Lynn	GP or practice nurse referral	Type 2 diabetes	Education program	Not specifically stated
Life and health with diabetes [98,99]	Buckinghamshire, Bedfordshire, Berkshire, Hertfordshire, Oxfordshire, and Uxbridge.	GP referral	Type 2 diabetes	Education program	Improve understanding of diabetes, confidence in self-management and quality of life
Living with diabetes [100]	Bristol	GP refers patient to program	Newly diagnosed type 2 diabetes (within last 12 months)	Education program	Healthier lifestyle, education

Name of project or service	Location	Referral process	Types of patients	Type of service, what is offered and by whom	Aim of service
Newham Community Prescriptions [101,102]	London (Newham)	Referral by GP to community prescription navigators	Type 2 diabetes and patients at risk	Personalized information and support service by community prescription navigators, physical activities offered by VSCE partners, gardening	Adherence to physical activity
Social prescribing project[a] [103,104]	Cullompton, Devon	GP refers patients to health facilitator who provides advice on exercise, diet, and so forth	Patients with CVD, type 2 diabetes, prediabetes, and other diseases	Individual advice, signposting service	Support patients to exercise and socialize
Start-Up exercise referral scheme[a] [105,106]	Cambridge	GP, practice nurse, or health professional refers to scheme	People with a range of medical conditions including diabetes	Exercise programs and support, assistance, and supervision from specialist exercise professional	Support patients whose health would benefit from leading a more active lifestyle

[a] Projects that are currently undergoing evaluation have been evaluated as part of a larger-scale evaluation or where the evaluation report was not available.

[b] GP: General practitioner.

Overview of Projects

The included projects represented a range of initiatives and interventions and were commissioned by a range of organizations, either solely or in collaboration, including NHS organizations, local government, and charities. Three projects had a national coverage, providing a standardized program (eg, DESMOND) [47] or a service that was locally adopted from an initial version (eg, X-PERT) [79] or a "social prescribing" intervention provided as part of a national program [24]. The remaining projects were limited to a local area, ranging from city districts to counties or regions.

Locating the Relevant Information

Although all published studies were located using a standard medical database, information on the projects or services was found from a variety of sources. Much of the information was retrieved from websites of the organizations delivering the projects. However, in order to include information on the prespecified aspects and, where evaluation reports were not available, it was necessary to follow links to other resources or websites and to access presentations, minutes of relevant meetings, and reports. The specific sources used are referenced in Table 2.

The studies published in peer-reviewed journals met our broad definition of social prescribing as a referral process from primary care to an external provider of nonmedical services in the community. These studies were not, however, described as social prescribing projects. In fact, only 1 study was located on PubMed that referred to social prescribing [107]. It was potentially relevant, as it was focused on long-term conditions. However, diabetes was not mentioned, and therefore it was not included. The descriptions varied considerably across the projects. Some were referred to as educational programs emphasizing the type of intervention, as "diabetes programs" emphasizing the condition, or GP referral programs focusing on the process. Projects or services were also described as a "community prescription," a community service, a voluntary, community, and social enterprise undertaking, or a social prescribing project.

The Range of Interventions

The intervention in the majority of projects (n=16) was a structured group education program that included multiple thematic areas such as disease information, disease management, healthy lifestyle, and health consequences of diabetes. For instance, the DESMOND program is a structured education program over 6 hours on 1 or 2 days [47]. In the X-PERT diabetes program, participants receive a handbook and are taught by trained educators about diabetes, weight management and other lifestyle factors, and diabetic complications [77]. A number of services involved referral or recommendation of exercise; 7 schemes were based solely on exercise ("Exercise on Prescription" schemes), whereas others offered access to exercise classes as part of a "menu" of community activities. For example, 1 service offered free physical activity options (eg, Tai Chi, Zumba, Community Gym) for 12 weeks at maximum or an individually tailored exercise program for 12 weeks in combination with personal support, a final meeting, follow-up contacts, and a report that was sent to the GP [101].

Of the interventions that did not involve group education or exercise on prescription, 1 involved trained and quality-assured lifestyle facilitators delivering an "individually tailored behavior change intervention" [22]. Another provided individual support to black and minority ethnic people with type 2 diabetes via "Public Health Improvement Teams" [84]. Other interventions included community-based diabetes prevention programs [48,50,60], a lifestyle and weight management program [76], and a social and horticultural project [64]. The remainder were described as "social prescribing." In one of these, a personalized signposting service was provided by "Social Prescribing Coordinators" to local services provided by a voluntary, community, and social enterprise (VSCE) group [30]. Similar services were provided by Community Health Champions, Community Prescription Navigators, Social Prescribers or Social Prescribing Workers (or VCS Advisors), link workers, and a practice-based health facilitator (or health manager).

Evidence from the Evaluations

We identified a variety of types of research and evaluation on the interventions, ranging from local projects with regular audits using questionnaires, to national programs tested using multicenter randomized controlled trials (RCTs). Of the 24 evaluations, 10 involved RCTs of behavioral, education, or exercise interventions. The majority were service evaluations that used a combination of approaches, both qualitative and quantitative. This required collection and analysis of monitoring data, measurement of outcomes using validated measures or self-reporting, cost-effectiveness calculations, and methods for gaining an insight into the perspectives of both providers and users of the services through surveys, interviews, and focus groups.

Of the education programs, the DESMOND program, one of the most extensively tested, had undergone a series of clinical trials and other evaluations. These include a large-scale RCT with a subsequent 3-year follow-up study [36,44], a more recent nonrandomized trial comparing health care professional and lay educators [45], and a study on cost-effectiveness [42]. The results revealed that its group education program for patients with newly diagnosed type 2 diabetes was effective at improving some key outcomes but not all. There were positive improvements in illness beliefs, which were still sustained after 3 years, but differences in biomedical or lifestyle outcomes were not apparent after this period [44]. Further research revealed that there was no difference in study outcomes between patients who were referred to 2 health professional educators and those who were referred to a team consisting of a professional educator and a layperson [45]. The cost-effectiveness analysis showed DESMOND to be cost-effective compared with usual care at 12 months [42]. Further studies focused on the facilitators and their skills, and the perspectives of participants in the program, through the use of qualitative approaches [39,46].

A second education-based intervention, X-PERT, has also been tested in an RCT, and a range of benefits were reported on physiological and behavioral outcomes, including improved glycemic control, BMI, diet, and self-management skills at 14 months [77]. A noncontrolled trial of an adaptation of the program designed for Bangladeshi adults with type 2 diabetes, which did not involve GP referral and therefore is not included in the tables, has also been completed but was underpowered due to low attendance rates [108,109]. A program designed specifically for Muslim diabetics was also assessed by retrospective analysis and reported beneficial effects [65].

Pre-post questionnaires and physiological measures were used to assess the CODE (Community Oriented Diabetes Education) program with positive trends but not significant changes in physiological measures. However, there was a significant increase in participants' knowledge scores, coping ability, motivation to change, and making informed decisions about their diabetes [33].

Other studies investigated the effectiveness of exercise on prescription schemes [26,54,71,75]. Mixed results were reported with effects, if observed, either limited to reduction of anxiety or not maintained over time. One RCT found no difference between an exercise program, instructor-led walking, or advice only [54], while a regression analysis of one scheme explored how the effects of support differed as a function of who provided the support [27]. RCTs were also used to assess the effects of interventions described as behavioral or lifestyle change interventions [22,51,60].

Internal or external evaluation of services described as social prescribing, either ongoing or completed, was reported on a number of the websites. Of these, evaluation reports were found for 6 "social prescribing" services. In other cases, the results of the evaluations were reported in brief on the website, or in the minutes of relevant meetings or other documents. Some services had been evaluated as part of large projects, whereas in several cases, reports were unavailable or could not be located. One further service had been approved by QISMET (the Quality Institute for Self-Management Education and Training), an independent body developed to support self-management providers and commissioners [98]. In some cases, the collection of data was reported but the results were not located by the searches we performed.

The type of evaluation ranged in design. Basic numerical data were collected including numbers referred to and accessing the service, and numbers of people reporting changes in various aspects such as increased physical activity or simply completing the program. Questionnaires and more complex designs, for example, mixed approaches using questionnaires combined with interviews or objective outcome data, were also employed in some cases.

One social prescribing project in inner London that underwent a full evaluation using a mixture of qualitative and quantitative methods also incorporated a matched control group [29]. No significant differences were found between groups based on the quantitative measures, although the qualitative data indicated beneficial effects with positive responses from the participants.

In the case of the other projects described as social prescribing for which full evaluations were found, a pre-post design was used [24,59,66,73]. Although positive outcomes were reported in, for example, physical activity, weight, blood pressure, mood, and social outcomes, (see Table 1 for the range of outcomes reported), this design did not allow firm conclusions on effectiveness to be drawn. One evaluation reported reduced hospital admissions and improved social outcomes [66]. Several evaluations focused on aspects related to implementing such a project in practice by exploring the perspectives of the patients, the health care providers, the role of the link worker, and the project governance to conclude with recommendations for future social prescribing projects [59].

While a formal, in-depth systematic appraisal of the methodology was not carried out, a number of issues were encountered when extracting data from the evaluations. It proved difficult in some reports to find the rationale, or occasionally the actual numbers, for the samples used in evaluations. It was also clear that missing data was an issue, particularly where follow-up was at several time points. This is, perhaps, not surprising considering that the aim of social prescribing and related interventions is to direct the patient to the community and voluntary sector rather than encourage follow-up in primary care. However, what was not always clear was how missing data were treated. Finally, as described earlier, referral was often based on psychological, social, or practical needs; this meant that the primary diagnosis was not often reported nor were the results provided by diagnosis.

Discussion

Principal Findings

Our study revealed limited evidence on social prescribing, specifically for type 2 diabetes in the United Kingdom and Ireland in peer-reviewed literature. Only 1 published study referred to social prescribing and that involved people with long-term conditions with no specific mention of type 2 diabetes. Nevertheless, broader Web-based searches demonstrated the existence of numerous “real-world” projects that investigated potentially relevant community-based interventions and had undergone evaluation using a variety of methods. Analysis of these projects also revealed the diverse ways in which the concept of social prescribing has been implemented in practice. The findings of this scoping study do not allow firm conclusions as to the effectiveness of social prescribing for people with type 2 diabetes. The results of the broader searches provide a clearer picture of the potential effects, possible outcomes, and processes involved, and thus inform research in this area. These evaluations also have the potential to provide valuable information for those organizations planning to implement similar services in future. This study has also revealed the different ways in which social prescribing may be described in practice, and this will aid those searching for evidence in future.

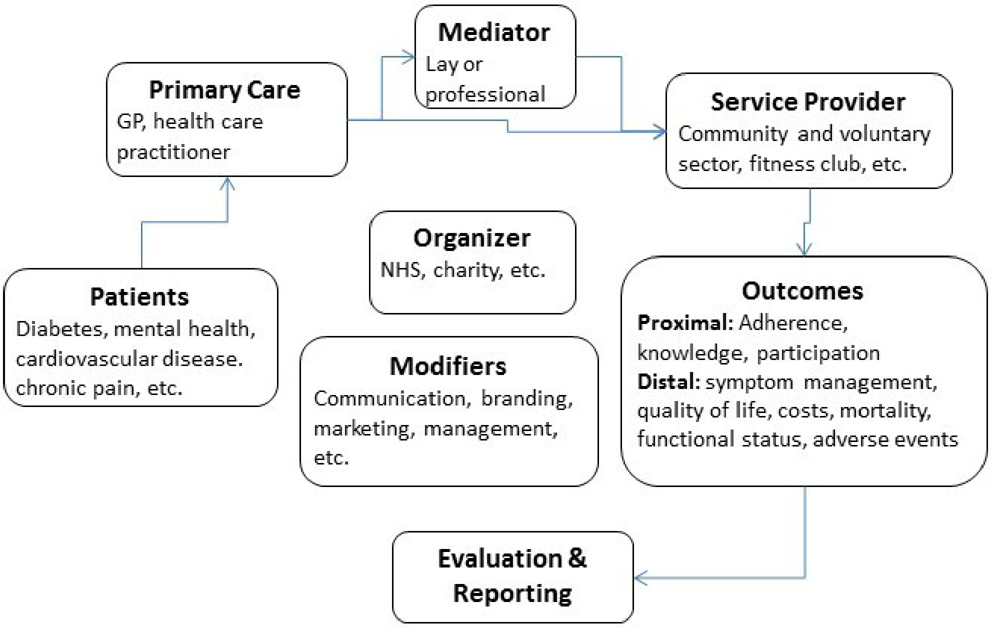
Challenges and Limitations

The challenges encountered when designing and conducting this study have led to a number of limitations. First, we attempted to identify evaluations of what is effectively a process, the referral of patients from primary care to a nonmedical, community-based activity, program, or service, rather than an intervention per se. In addition, some studies may not have referred specifically to social prescribing and described the intervention using different wording due to the lack of consensus on terminology in this area, thus precluding their identification. We attempted to address these points by using a range of several search strategies and an iterative process whereby we amended the later strategies according to the results of the initial searches. Although we extended the later database searches by including terms for prediabetes and a range of interventions, because of the recent proliferation of websites on the prevention of diabetes, we did not search specifically for diabetes prevention programs on the Web. Such programs, for example, those launched as part of the national Diabetes Prevention Program [110], are likely to use similar models and interventions to the projects included in this review, but further exploration of this aspect would be valuable. The selection of studies also relied on the referral process being reported clearly. The complexity of both identifying and assessing the evidence on social prescribing is demonstrated by the case of the X-PERT project. The RCT on X-PERT was initially excluded from our search because it did not include referral from primary care to an external service provider, whereas in practice, patients are referred from the general practice to the X-PERT educators via social prescribing. A similar situation occurred with later trials of the DESMOND program. Related to this point, it is debatable whether referring patients to a program as part of an RCT is adequately replicating the social prescribing process in practice. A conventional systematic review of published literature might, therefore, not represent a feasible approach to identify the entire literature and assess the effectiveness of the process.

A second potential limitation was that we limited the search to the United Kingdom and Ireland. The heterogeneity of national health systems, the plethora of diabetes websites, and the varied language used in describing this concept outside the United Kingdom would have proved a considerable challenge, had we attempted to cover the worldwide literature. This does mean that broadly similar projects in other countries were omitted from this review. Third, only information available on the Internet was retrieved and there was no direct contact with those delivering programs or services for further information, confirmation, or clarification. For this reason, it is possible we might have omitted relevant projects or incorrectly interpreted the information we were able to locate. Website information also evolved throughout the course of the study. Finally, by focusing on diabetes and including only those programs or services that specifically referred to people with this condition, we might have omitted relevant services or programs. Conversely, using a broad definition of the intervention resulted in the inclusion of projects and services that, perhaps, do not fit with the original concept of social prescribing as being a mechanism to refer from primary care to the voluntary and community sector. In some cases, the services, although nonmedical and community-based, were still provided by health professionals.

Nevertheless, our broad searches have demonstrated that social prescribing is prevalent in the United Kingdom, and that information is available that is potentially valuable but not currently in the published literature and therefore difficult to locate and access. At this point, our analysis on the range of projects and services, and the evaluations of these, may prove helpful.

Figure 2. A model of social prescribing including people with type 2 diabetes.



The Social Prescribing Concept and Terminology Surrounding This Concept

It is clear from the findings of this review that different models have been implemented in practice, some of which fit more closely with the original concept of "social prescribing." It is also clear that although the term "social prescribing" is more widely used in practice in the United Kingdom, it is not a concept that is currently recognized in the main medical databases or in other countries. Even within the United Kingdom, the link workers, who are a crucial part of many social prescribing services, are described in different ways. The findings from this scoping review do, however, suggest a preliminary model for social prescribing such as that presented in Figure 2. This will require further development and refinement.

Translating Research Into Practice

It has been argued that the term "translational research" does not distinguish between testing of new treatments and research on how to implement new treatments in practice [10]. This is supported by our finding that the process of social prescribing, potentially the crucial step in the implementation process, is neglected in comparison to the intervention itself. A wide range of models of the process were revealed, for example, direct referral by the GP to the community sector, referral to the community via an in-practice link worker, or referral to a link worker based in the voluntary and community sector. Only 1 study attempted to address whether the process of prescribing an intervention had any influence by comparing referral to exercise programs in the community with tailored advice and information on local exercise facilities. None of the evaluations we located investigated, for example, how direct referral by a GP compared with referral via a link worker.

In general, many barriers exist that hamper the adoption of research into diabetes into clinical care at the level of the patient, primary care, health care system, community, and society [111]. Several countries have developed diabetes prevention programs for clinical care [112], whereas there has been limited translational research in the United Kingdom [113]. The effects of lifestyle programs in real-world primary care for patients with prediabetes or T2DM are small [114] and depend substantially on the program and its implementation [115]. Development, evaluation, and reporting of translational research programs need to be adapted to reflect the context and consider generalizability [116]. In exploring the effectiveness of implementation, various measures of success need to be taken into account and interpreted. For example, although in many cases, qualitative data in the evaluations found in our searches revealed a positive patient experience, this has not yet been supported by quantitative outcomes to the same degree. It is also clear that, in evaluating the effectiveness of these programs, a range of generic outcomes are used and the focus is often not on disease-specific indicators such as hemoglobin A1c. This is because social prescribing grew out of a perceived need to address broader social, behavioral, and practical issues that impact on, and may be caused by, the disease. Thus, the focus is on the person who has diabetes rather than on the disease itself. Interest in social prescribing is increasing, and its effects on specific groups such as those with diabetes are likely to be of interest, making it important to provide guidance on how real-world evaluations should be conducted and reported to ensure that the data collected can be more accurately compared and analyzed. The results of our analysis of real-world project evaluation will inform this guidance, while guidance such as that produced by Diabetes UK on commissioning to involve nontraditional providers in the support of people with long-term conditions is also valuable [117].

Our review also suggests a need to improve the reporting of implementation research, specifically service evaluations. Social, cultural, legal, demographic, and other factors that are controlled in trials become relevant influences during the process of implementation [118]. Combining these factors with the diversity in settings, methods, and outcomes of the real-world projects indicates that further development of existing reporting guidelines would be beneficial [119]. However, although standardization of reporting of evaluations would allow greater comparison across services and projects, a service evaluation is produced for a specific organization and to meet the specific requirements of that organization. Thus, increasing standardization of reporting, should this be considered feasible, must be balanced with the need to meet these requirements.

It has been suggested that evidence synthesis plays a pivotal role in developing or refining a framework of translational research [120]. A first step to enable this would be to better understand and define the multiple dimensions of translational research including implementation projects in real-world settings [10]. This might begin with a debate on reporting of different types of translational research, such has been proposed for Web-based health interventions [121]. Additionally, the challenges encountered in identifying relevant projects in this study indicate that it might prove fruitful to develop specific databases or networks to support larger-scale and more rigorous evaluation and comparison of programs and their relative cost-effectiveness.

Finally, a discussion might be necessary on how the principles of systematic reviews of health interventions can be applied to translational research, as the hierarchical model of evidence-based medicine has been shown valid only for certain questions such as efficacy [122]. It has been suggested that the study design alone may be an inadequate marker of the quality of evidence for interventions that are complex, program based, and dependent on context [123]. Consideration must be given to the evaluation process and its ability to detect effects and distinguish between failure of an intervention and failure of the delivery process. It is clear from this review that approaches incorporating a range of methods are necessary for evaluating the implementation of such complex interventions in practice. Novel approaches to synthesizing the evidence in this field are already underway [124].

As part of this discussion, it will be worth considering what role service evaluations of real-world projects play in the evidence base, and whether there are minimum standards for conducting such evaluations. Carrying out formal quality appraisal of these is complex, as they involve a range of data collection approaches, and it is unclear how best to accurately assess and compare the methodology used in individual evaluations. Although standardized tools and checklists exist for designs such as RCTs and qualitative studies, similar tools are not widely available for complex designs. As described earlier, there is also a question of how service evaluations may best be identified, as it is clear that the majority remain in the gray literature and would not be located via conventional systematic search techniques.

Conclusions

This review revealed the range of models of social prescribing and related nonmedical community-based services, and the extent of evaluation that has been carried out to assess the potential benefits of these, which would not have been achieved by searching databases alone. Although the evidence from these evaluations does not prove that social prescribing is an effective measure for implementing nonmedical interventions for patients with type 2 diabetes in the United Kingdom, there is sufficient to indicate the value of further evaluation and comparison, particularly if focused on real-world settings. The findings of this scoping study may inform future evaluation. Further research is necessary to better understand the communication between primary care and community or voluntary sector and to improve the documentation of the final step of translational research, implementing interventions in real-world settings.

Accessing Web-based information provides a potential method for investigating how specific innovative health concepts, such as social prescribing for type 2 diabetes, have been implemented in practice and the full extent of the evaluation of such innovations. Several challenges were encountered including defining the concept, focusing on process plus intervention, and systematically searching varied and evolving Web-based sources. Obtaining sufficient relevant information requires searching for and analyzing information from a range of sources. The methods and findings from this study have already informed a broader scoping exercise on the evaluation of social prescribing projects in the United Kingdom, which is underway. Further exploration of this approach will inform future research on the application of health-related concepts into practice.

Acknowledgments

This study was partially funded by the Make My Day Better charity as part of an evaluation of a social prescribing service. The charity had no involvement in the preparation, review, or approval of this paper.

Authors' Contributions

KP contributed to the design of the study, conducted the follow-up searches, selected papers and websites, extracted data, and contributed to the writing of the manuscript. ML contributed to the design of the study, conducted the initial and iterative searches, selected papers and websites, extracted data, and contributed to the writing of the manuscript. MP contributed to the design of the study, checked the selections and data extraction, and contributed to the writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategies.

[PDF File (Adobe PDF File), 27KB - jmir_v19i2e20_app1.pdf]

References

1. World Health Organisation (WHO). 2016. Diabetes: Fact sheet URL: http://www.who.int/mediacentre/factsheets/fs312/en [accessed 2016-07-05] [WebCite Cache ID 6imUcb5zZ]
2. Holman N, Forouhi NG, Goyder E, Wild SH. The association of public health observatories (APHO) diabetes prevalence model: estimates of total diabetes prevalence for England, 2010-2030. Diabet Med 2011 May;28(5):575-582. [doi: 10.1111/j.1464-5491.2010.03216.x] [Medline: 21480968]
3. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
4. Schellenberg ES, Dryden DM, Vandermeer B, Ha C, Korownyk C. Lifestyle interventions for patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. Ann Intern Med 2013 Oct 15;159(8):543-551. [doi: 10.7326/0003-4819-159-8-201310150-00007] [Medline: 24126648]
5. Stark CS, Fradkin JE, Saydah SH, Rust KF, Cowie CC. The prevalence of meeting A1C, blood pressure, and LDL goals among people with diabetes, 1988-2010. Diabetes Care 2013 Aug;36(8):2271-2279 [FREE Full text] [doi: 10.2337/dc12-2258] [Medline: 23418368]
6. Picker Institute Europe. 2007. National survey of local health services 2006 URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/StatisticalWorkAreas/Statisticalhealthcare/DH_073494 [accessed 2017-01-03] [WebCite Cache ID 6nEuPvWWS]
7. NHS England. 2014. Five year forward view URL: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf [accessed 2017-01-03] [WebCite Cache ID 6nEuYE1pT]
8. South J, Higgins TJ, Woodall J, White SM. Can social prescribing provide the missing link? Prim Health Care Res Dev 2008 Nov 4;9(04):310. [doi: 10.1017/S146342360800087X]
9. Brandling J, House W. Social prescribing in general practice: adding meaning to medicine. Br J Gen Pract 2009 Jun;59(563):454-456 [FREE Full text] [doi: 10.3399/bjgp09X421085] [Medline: 19520038]
10. Woolf SH. The meaning of translational research and why it matters. J Am Med Assoc 2008 Jan 9;299(2):211-213. [doi: 10.1001/jama.2007.26] [Medline: 18182604]
11. Higgins J, Green S. 2011. Cochrane handbook for systematic reviews of interventions, Version 5 URL: http://handbook.cochrane.org/ [WebCite Cache ID 6lnGIdQud]
12. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
13. Centre for Reviews and Dissemination. University of York; 2015. Evidence to inform the commissioning of social prescribing URL: https://www.york.ac.uk/media/crd/Ev%20briefing_social_prescribing.pdf [accessed 2016-11-05] [WebCite Cache ID 6lnGvuApn]
14. Thomson L, Camic PM, Chatterjee H. 2015. Social prescribing: a review of community referral schemes URL: https://www.ucl.ac.uk/museums/research/museumsonprescription [accessed 2017-01-03] [WebCite Cache ID 6nEvO4m5T]
15. Campbell F, Holmes M, Everson-Hock E, Davis S, Buckley WH, Anokye N, et al. A systematic review and economic evaluation of exercise referral schemes in primary care: a short report. Health Technol Assess 2015 Jul;19(60):1-110 [FREE Full text] [doi: 10.3310/hta19600] [Medline: 26222987]
16. Murphy S, Raisane L, Moore G, Edwards R, Linck P, Hounsome N, et al. 2010. The evaluation of the national exercise referral scheme in Wales URL: http://gov.wales/docs/caecd/research/101104nationalexerciseschemeen.pdf [accessed 2016-11-06] [WebCite Cache ID 6lnf1rnDC]
17. Mental Health Foundation. 2009. Moving on Up URL: https://www.mentalhealth.org.uk/sites/default/files/moving_on_up_report.pdf [accessed 2016-11-06] [WebCite Cache ID 6lnf85ln7]
18. Year of Care. 2011. Year of Care: Report of findings from the pilot programme URL: http://www.yearofcare.co.uk/sites/default/files/images/YOC_Report%20-%20correct.pdf [accessed 2016-11-06] [WebCite Cache ID 6lnnd06NE]
19. Ecorys UK and University of Keele. Fit as a Fiddle Final Evaluation Report URL: http://www.ageuk.org.uk/Documents/EN-GB/ID201168_Fit_As_A_Fiddle_Evaluation_Report_FINAL130313_FINAL.pdf?pdf?true?dtrk=true [accessed 2016-11-06] [WebCite Cache ID 6lnnfI3IP]
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res Synth Methods 2014 Dec;5(4):371-385 [FREE Full text] [doi: 10.1002/jrsm.1123] [Medline: 26052958]

22. Griffin SJ, Simmons RK, Prevost AT, Williams KM, Hardeman W, Sutton S, et al. Multiple behaviour change intervention and outcomes in recently diagnosed type 2 diabetes: the ADDITION-Plus randomised controlled trial. Diabetologia 2014 Jul;57(7):1308-1319 [FREE Full text] [doi: 10.1007/s00125-014-3236-6] [Medline: 24759957]
23. Dambha-Miller H, Cooper AJ, Simmons RK, Kinmonth AL, Griffin SJ. Patient-centred care, health behaviours and cardiovascular risk factor levels in people with recently diagnosed type 2 diabetes: 5-year follow-up of the ADDITION-Plus trial cohort. BMJ Open 2016 Jan 06;6(1):e008931 [FREE Full text] [doi: 10.1136/bmjopen-2015-008931] [Medline: 26739725]
24. Wigfield A, Alden S, Kispeter E. 2015. Age UK's fit for the future Social Prescribing extension project Evaluation report URL: http://www.ageuk.org.uk/Documents/EN-GB/fit%20as%20a%20fiddle/fftf%20extension%20project-Final%20Evaluation%20Report%20(July%202015).docx.pdf?dtrk=true [accessed 2016-11-28] [WebCite Cache ID 6mMCWqI04]
25. Wigfield A, Kispeter E, Alden S, Turner R. 2015. Age UK's Fit for the Future project: Evaluation report URL: http://www.ageuk.org.uk/Documents/EN-GB/fit%20as%20a%20fiddle/fit%20for%20the%20future%20project%20-%20Final%20Evaluation%20Report%20(July%202015).pdf?dtrk=true [accessed 2016-11-28] [WebCite Cache ID 6mMCU27D4]
26. Duda JL, Williams GC, Ntoumanis N, Daley A, Eves FF, Mutrie N, et al. Effects of a standard provision versus an autonomy supportive exercise referral programme on physical activity, quality of life and well-being indicators: a cluster randomised controlled trial. Int J Behav Nutr Phys Act 2014;11:10 [FREE Full text] [doi: 10.1186/1479-5868-11-10] [Medline: 24475766]
27. Rouse PC, Ntoumanis N, Duda JL, Jolly K, Williams GC. In the beginning: role of autonomy support on the motivation, mental health and intentions of participants entering an exercise referral scheme. Psychol Health 2011 Jun;26(6):729-749. [doi: 10.1080/08870446.2010.492454] [Medline: 21827332]
28. Birmingham City Council. About Excercise (sic) on Prescription URL: http://www.birmingham.gov.uk/cs/Satellite?c=Page&childpagename=Exercise-on-Prescription%2FPageLayout&cid=1223237047303&pagename=BCC%2FCommon%2FWrapper%2FWrapper [WebCite Cache ID 6imVlAx8H]
29. City and Hackney Clinical Commissioning Group, University of East London. 2014. Shine 2014 final report: Social Prescribing: integrating GP and Community Assets for Health URL: http://www.health.org.uk/sites/health/files/City%20and%20Hackney%20CCG%20final%20report.pdf [accessed 2017-01-04] [WebCite Cache ID 6nGIomsAQ]
30. City and Hackney Clinical Commissioning Group. About Us URL: http://www.cityandhackneyccg.nhs.uk/about-us/ [accessed 2016-07-06] [WebCite Cache ID 6imVyOW2F]
31. City and Hackney Clinical Commissioning Group. 2014. Building Health Partnerships: Social prescribing in City and Hackney URL: http://www.cityandhackneyccg.nhs.uk/news/Building-Health-Partnerships-Social-prescribing-in-City-and-Hackney.htm [accessed 2016-07-06] [WebCite Cache ID 6imW39Sp6]
32. University of East London. Social prescribing: integrating GP and community assets for health URL: https://www.uel.ac.uk/IHHD/Our-work/Research-themes-and-projects/Well-communities-temp/Health-Foundation-Shrine [accessed 2016-07-06] [WebCite Cache ID 6imWInAmR]
33. Clarke A. Effectiveness of a community orientated diabetes education (CODE) programme for people with type 2 diabetes. European Diabetes Nursing 2015 Feb 17;8(3):94-99a. [doi: 10.1002/edn.188]
34. Diabetes Federation of Ireland. 2008. Community Orientated Diabetes Education (CODE): A structured education programme for people with diabetes. Evaluation Report 2007 URL: https://www.diabetes.ie/wp-content/uploads/2014/11/CODE_2007_Report_April_2008_rev.pdf [accessed 2017-01-03] [WebCite Cache ID 6nF3Bh230]
35. Diabetes Ireland. CODE: Community Orientated Diabetes Education URL: https://www.indi.ie/images/public_docs/2010_CODE_info_for_INDIx.pdf [accessed 2016-07-27] [WebCite Cache ID 6jJcvG4kU]
36. Davies MJ, Heller S, Skinner TC, Campbell MJ, Carey ME, Cradock S, Diabetes Education and Self Management for Ongoing and Newly Diagnosed Collaborative. Effectiveness of the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cluster randomised controlled trial. Br Med J 2008 Mar 1;336(7642):491-495 [FREE Full text] [doi: 10.1136/bmj.39474.922025.BE] [Medline: 18276664]
37. Skinner TC, Carey ME, Cradock S, Daly H, Davies MJ, Doherty Y, et al. Diabetes Education and Self-Management for Ongoing and Newly Diagnosed (DESMOND): process modelling of pilot study. Patient Educ Couns 2006 Dec;64(1-3):369-377. [doi: 10.1016/j.pec.2006.04.007] [Medline: 17011154]
38. Khunti K, Skinner TC, Heller S, Carey ME, Dallosso HM, Davies MJ. Biomedical, lifestyle and psychosocial characteristics of people newly diagnosed with type 2 diabetes: baseline data from the DESMOND randomized controlled trial. Diabet Med 2008 Dec;25(12):1454-1461. [doi: 10.1111/j.1464-5491.2008.02620.x] [Medline: 19046245]
39. Ockleford E, Shaw RL, Willars J, Dixon-Woods M. Education and self-management for people newly diagnosed with type 2 diabetes: a qualitative study of patients' views. Chronic Illn 2008 Mar;4(1):28-37. [doi: 10.1177/1742395307086673] [Medline: 18322027]

40. Skinner TC, Carey ME, Cradock S, Dallosso HM, Daly H, Davies MJ, et al. 'Educator talk' and patient change: some insights from the DESMOND (Diabetes Education and Self Management for Ongoing and Newly Diagnosed) randomized controlled trial. Diabet Med 2008 Sep;25(9):1117-1120. [doi: 10.1111/j.1464-5491.2008.02492.x] [Medline: 19183318]
41. Skinner TC, Carey ME, Cradock S, Dallosso HM, Daly H, Davies MJ, DESMOND Collaborative. Depressive symptoms in the first year from diagnosis of Type 2 diabetes: results from the DESMOND trial. Diabet Med 2010 Aug;27(8):965-967. [doi: 10.1111/j.1464-5491.2010.03028.x] [Medline: 20653757]
42. Gillett M, Dallosso HM, Dixon S, Brennan A, Carey ME, Campbell MJ, et al. Delivering the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cost effectiveness analysis. Br Med J 2010;341:c4093 [FREE Full text] [Medline: 20729270]
43. Skinner TC, Carey ME, Cradock S, Dallosso HM, Daly H, Davies MJ, DESMOND Collaborative. Comparison of illness representations dimensions and illness representation clusters in predicting outcomes in the first year following diagnosis of type 2 diabetes: results from the DESMOND trial. Psychol Health 2011 Mar;26(3):321-335. [doi: 10.1080/08870440903411039] [Medline: 20309775]
44. Khunti K, Gray LJ, Skinner T, Carey ME, Realf K, Dallosso H, et al. Effectiveness of a diabetes education and self management programme (DESMOND) for people with newly diagnosed type 2 diabetes mellitus: three year follow-up of a cluster randomised controlled trial in primary care. Br Med J 2012;344:e2333 [FREE Full text] [Medline: 22539172]
45. Carey ME, Mandalia PK, Daly H, Gray LJ, Hale R, Martin SL, et al. Increasing capacity to deliver diabetes self-management education: results of the DESMOND lay educator non-randomized controlled equivalence trial. Diabet Med 2014 Nov;31(11):1431-1438. [doi: 10.1111/dme.12483] [Medline: 24798205]
46. Eborall HC, Dallosso HM, Daly H, Martin-Stacey L, Heller SR. The face of equipoise--delivering a structured education programme within a randomized controlled trial: qualitative study. Trials 2014;15:15 [FREE Full text] [doi: 10.1186/1745-6215-15-15] [Medline: 24405854]
47. Diabetes Education and Self Management for Ongoing and Newly Diagnosed (DESMOND) project URL: http://www.desmond-project.org.uk/index.php [accessed 2016-07-27] [WebCite Cache ID 6jJdnNvJn]
48. Davies MJ, Gray LJ, Troughton J, Gray A, Tuomilehto J, Farooqi A, Let's Prevent Diabetes Team. A community based primary prevention programme for type 2 diabetes integrating identification and lifestyle intervention for prevention: the Let's Prevent Diabetes cluster randomised controlled trial. Prev Med 2016 Mar;84:48-56. [doi: 10.1016/j.ypmed.2015.12.012] [Medline: 26740346]
49. Gray LJ, Yates T, Troughton J, Khunti K, Davies MJ, Let’s Prevent Diabetes Team. Engagement, retention, and progression to type 2 diabetes: a retrospective analysis of the cluster-randomised “Let's Prevent Diabetes” trial. PLoS Med 2016 Jul;13(7):e1002078 [FREE Full text] [doi: 10.1371/journal.pmed.1002078] [Medline: 27404094]
50. Yates T, Edwardson CL, Henson J, Gray LJ, Ashra NB, Troughton J, et al. Walking away from type 2 diabetes: a cluster randomized controlled trial. Diabet Med 2016 Sep 2:- Epub ahead of print. [doi: 10.1111/dme.13254] [Medline: 27589017]
51. Penn L, White M, Oldroyd J, Walker M, Alberti KG, Mathers JC. Prevention of type 2 diabetes in adults with impaired glucose tolerance: the European Diabetes Prevention RCT in Newcastle upon Tyne, UK. BMC Public Health 2009 Sep 16;9:342 [FREE Full text] [doi: 10.1186/1471-2458-9-342] [Medline: 19758428]
52. Penn L, Moffatt SM, White M. Participants' perspective on maintaining behaviour change: a qualitative study within the European Diabetes Prevention Study. BMC Public Health 2008 Jul 10;8:235 [FREE Full text] [doi: 10.1186/1471-2458-8-235] [Medline: 18616797]
53. Penn L, White M, Lindström J, den Boer AT, Blaak E, Eriksson JG, et al. Importance of weight loss maintenance and risk prediction in the prevention of type 2 diabetes: analysis of European Diabetes Prevention Study RCT. PLoS One 2013;8(2):e57143 [FREE Full text] [doi: 10.1371/journal.pone.0057143] [Medline: 23451166]
54. Isaacs AJ, Critchley JA, Tai SS, Buckingham K, Westley D, Harridge SD, et al. Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. Health Technol Assess 2007 Mar;11(10):1-165, iii-iv [FREE Full text] [Medline: 17313906]
55. Solomon E, Clifford Z. 2015. Evaluation of the “Living Well, Taking Control” community-based prevention and management of type 2 diabetes programme: a study protocol and presentation of preliminary findings URL: http://decipher.uk.net/wp-content/uploads/2015/03/F2-Emma-Solomon-Diabetes-prevention.pdf [accessed 2016-11-06] [WebCite Cache ID 6lovTbfxj]
56. Health Exchange. 2016. Living Well, Taking Control URL: https://www.healthexchange.org.uk/corporate/our-partners/living-well-taking-control/ [accessed 2016-11-28] [WebCite Cache ID 6mMGWNnES]
57. Jennings E, Greene P, Nasteva G, Mellor D, Cowley P, Harvey J. Evaluation of Initial Structured Group Education in Newly Diagnosed Type 2 Diabetes URL: http://connect.qualityincare.org/__data/assets/pdf_file/0009/547074/Evaluation_of_Initial_Structured_Group_Education_in_Newly_Diagnosed_Type_2_Diabetes.pdf [WebCite Cache ID 6lokZAtzn]
58. North Wales NHS Trust. Newly Diagnosed Type 2 Diabetes Dietary Education Group (Type 2 group leaflet) URL: http://connect.qualityincare.org/__data/assets/pdf_file/0011/547076/Type_2_Group_Leaflet.pdf [accessed 2017-01-03] [WebCite Cache ID 6nF7DSP8C]

59. HeathWORKS Newcastle. 2013. Newcastle Social Prescribing Project: Final report URL: http://www.healthworksnewcastle.org.uk/2013/08/newcastle-social-prescribing-project-final-evaluation/ [accessed 2017-01-03] [WebCite Cache ID 6nF7sJKAJ]
60. Penn L, Ryan V, White M. Feasibility, acceptability and outcomes at a 12-month follow-up of a novel community-based intervention to prevent type 2 diabetes in adults at high risk: mixed methods pilot study. BMJ Open 2013 Nov 12;3(11):e003585 [FREE Full text] [doi: 10.1136/bmjopen-2013-003585] [Medline: 24227871]
61. Penn L, Dombrowski SU, Sniehotta FF, White M. Participants' perspectives on making and maintaining behavioural changes in a lifestyle intervention for type 2 diabetes prevention: a qualitative study using the theory domain framework. BMJ Open 2013 Jun 28;3(6):pii: e002949 [FREE Full text] [doi: 10.1136/bmjopen-2013-002949] [Medline: 23811173]
62. Penn L, Dombrowski SU, Sniehotta FF, White M. Perspectives of UK Pakistani women on their behaviour change to prevent type 2 diabetes: qualitative study using the theory domain framework. BMJ Open 2014 Jul 08;4(7):e004530 [FREE Full text] [doi: 10.1136/bmjopen-2013-004530] [Medline: 25005595]
63. Middlesbrough Environment City (Mevcity). New Life New You: Preventing Type 2 diabetes in BME communities in Middlesbrough URL: http://www.menvcity.org.uk/nlny/ [accessed 2016-11-06] [WebCite Cache ID 6lovYpRTr]
64. Clift S, Cunningham L, Winczewska N. PoLLeN People, Life, Landscape and Nature: An evaluation 2012-2013. Canterbury: Canterbury Christ Church University; 2013.
65. Bravis V, Hui E, Salih S, Mehar S, Hassanein M, Devendra D. Ramadan Education and Awareness in Diabetes (READ) programme for Muslims with type 2 diabetes who fast during Ramadan. Diabet Med 2010 Mar;27(3):327-331. [doi: 10.1111/j.1464-5491.2010.02948.x] [Medline: 20536496]
66. Dayson C, Bashir N. 2014. The social and economic impact of the Rotherham Social Prescribing Pilot Main Evaluation Report URL: http://www4.shu.ac.uk/research/cresr/sites/shu.ac.uk/files/social-economic-impact-rotherham.pdf [WebCite Cache ID 6mSK5LK2K]
67. Dayson C, Bashir N, Pearson S. 2013. From dependence to independenc: emerging lessons from the Rotherham Social Prescribing Pilot Summary Report URL: http://www4.shu.ac.uk/research/cresr/sites/shu.ac.uk/files/social-economic-impact-rotherham-summary.pdf [WebCite Cache ID 6mSKCFduv]
68. Voluntary Action Rotherham. Social Prescribing service URL: http://www.varotherham.org.uk/social-prescribing-service/ [accessed 2016-07-27] [WebCite Cache ID 6jJoNodmg]
69. Warwickshire Health and Wellbeing. Connect WELL: Social Prescribing Pilot Project in Rugby URL: http://hwb.warwickshire.gov.uk/2015/03/06/connect-well-social-prescribing-pilot-project-in-rugby/ [accessed 2017-01-04] [WebCite Cache ID 6nGJLPIfi]
70. White J, Coan S. 2016. Sadee Smile Evaluation Final report URL: http://www.touchstonesupport.org.uk/wp-content/uploads/2016/08/Sadee-Smile-Evaluation-Full-Report.pdf [accessed 2016-11-06] [WebCite Cache ID 6loviZUkc]
71. Flannery O, Loughren E, Baker C, Crone D. 2014. Exercise on Prescription Evaluation Report for South Gloucestershire URL: http://eprints.glos.ac.uk/2379/ [accessed 2016-11-06] [WebCite Cache ID 6lovnXwOX]
72. South Gloucestershire Council. Exercise on prescription URL: http://www.southglos.gov.uk/health-and-social-care/staying-healthy/getting-active/exercise-on-prescription/ [accessed 2016-07-27] [WebCite Cache ID 6jJeZpOWc]
73. Jones M, Kimberlee R, Deave T. 2009. South West Well-being Programme Evaluation Report URL: http://www.cles.org.uk/wp-content/uploads/2011/03/SWWB-Report-March-2009.pdf [accessed 2016-11-30] [WebCite Cache ID 6mPGTCP6U]
74. Kimberlee R. 2013. Developing a social prescribing approach for Bristol Project Report URL: http://eprints.uwe.ac.uk/23221/ [accessed 2016-07-27] [WebCite Cache ID 6jJpj469U]
75. Harrison RA, Roberts C, Elton PJ. Does primary care referral to an exercise programme increase physical activity one year later? A randomized controlled trial. J Public Health (Oxf) 2005 Mar;27(1):25-32 [FREE Full text] [doi: 10.1093/pubmed/fdh197] [Medline: 15564275]
76. NHS Wirral.: University of Liverpool Wirral PCT Lifestyle and Weight Management Programme: A preliminary report by the Health Economics Unit University of Liverpool Management School URL: http://info.wirral.nhs.uk/intelligencehub/healtheconomicsandvalueformoney.html [accessed 2017-01-04] [WebCite Cache ID 6nGKA39hT]
77. Deakin TA, Cade JE, Williams R, Greenwood DC. Structured patient education: the diabetes X-PERT Programme makes a difference. Diabet Med 2006 Sep;23(9):944-954. [doi: 10.1111/j.1464-5491.2006.01906.x] [Medline: 16922700]
78. X-PERT Health. Welcome to X-PERT Health: Diabetes Programme URL: http://www.xperthealth.org.uk/ [accessed 2017-01-04] [WebCite Cache ID 6nGKQil6b]
79. X-PERT Health. People with diabetes URL: http://www.xperthealth.org.uk/people-with-diabetes/x-pert-diabetes [accessed 2016-07-27] [WebCite Cache ID 6jJqdLzrc]
80. Altogether Better. Altogether Better Diabetes URL: http://www.altogetherbetter.org.uk/altogether-better-diabetes [accessed 2016-11-06] [WebCite Cache ID 6low0egsW]
81. Altogether Better. 2015. Altogether Better Working Together to Create Healthier People and Communities URL: http://www.altogetherbetter.org.uk/evidence-and-resources [accessed 2016-11-07] [WebCite Cache ID 6lpJSaHVR]
82. Birmingham City Council. Be active plus URL: http://www.birmingham.gov.uk/beactiveplus [accessed 2016-07-06] [WebCite Cache ID 6imVNM6Pp]

83. Be Active Birmingham. 2016. Information for surgeries URL: http://beactivebirmingham.co.uk/be-active-plus/information-for-surgeries [accessed 2016-07-06] [WebCite Cache ID 6imVcZKCr]
84. Building Health Partnerships--No. 8: Improving access to targeted services for BME diabetics in Bristol URL: http://www.socialenterprise.org.uk/public/uploads/editor/files/Health_Partnership_Bristol_08_final.pdf [accessed 2016-07-05] [WebCite Cache ID 6imVqLAoG]
85. Camden. 2016. Camden Exercise Referral URL: http://www.camden.gov.uk/ccm/content/leisure/sports-and-activities/activities-for-over-50s/active-health-classes.en [accessed 2016-07-06] [WebCite Cache ID 6imWoPHz8]
86. Camden and Islington Exercise Referral & Adult weight management service: Inclusion and Exclusion criteria URL: https://gps.camdenccg.nhs.uk/cdn/serve/service-downloads/1458746654-2b49340edaa91f835b545c5a92c75a4a.pdf [accessed 2017-01-04] [WebCite Cache ID 6nGKuaWfV]
87. Stathi A, Milton K, Riddoch C. 2006. Evaluation of the London Borough of Camden Exercise Referral Scheme URL: http://opus.bath.ac.uk/48825/ [accessed 2017-01-06] [WebCite Cache ID 6nJFlU9Vi]
88. Kent Community Health NHS Foundation Trust. Diabetes Education and Revision in East Kent (DEREK) URL: http://www.kentcht.nhs.uk/wp-content/uploads/2016/03/DEREK-programme-00408.pdf [accessed 2017-01-04] [WebCite Cache ID 6nGLD1qWZ]
89. Berkshire Healthcare NHS Foundation Trust. Diabetes URL: http://www.berkshirehealthcare.nhs.uk/page.asp?fldArea=1&fldMenu=12&fldSubMenu=0&fldKey=138 [WebCite Cache ID 6jJeNDbv7]
90. Positive Step. Positive step (Go4life): Supporting wellness in North Somerset URL: http://positivestep.org.uk/ [accessed 2016-07-27] [WebCite Cache ID 6jJepeGKv]
91. North Somerset Council. Go4life: Active directory and publications URL: http://www.n-somerset.gov.uk/my-services/leisure/sport/getting-active/go4life/ [accessed 2016-07-27] [WebCite Cache ID 6jJfY6pln]
92. North Somerset Pathways. Healthy Lifestyles URL: http://www.northsomersetpathways.co.uk/healthy_lifestyles.php [accessed 2017-01-04] [WebCite Cache ID 6nGLfzFnF]
93. Your Diabetes. Your diabetes. Structured education (Good2go) URL: http://www.yourdiabetes.org.uk/health_professionals/structured_education/ [accessed 2017-01-04] [WebCite Cache ID 6nGLpYfbA]
94. North Yorkshire County Council. Diabetes URL: http://www.northyorks.gov.uk/article/27407/Diabetes [accessed 2016-07-27] [WebCite Cache ID 6jJg2JdhW]
95. Harrogate and District NHS Foundation Trust. Diabetes service URL: https://www.hdft.nhs.uk/services/endocrinology/diabetes-service/ [accessed 2016-07-27] [WebCite Cache ID 6jJgG0XPZ]
96. Nottingham University Hospitals NHS Trust. Structured education for type 2 diabetes (Juggle) URL: http://www.nuh.nhs.uk/our-services/services/diabetes/clinics-and-services/structured-education-for-type-2-diabetes/ [accessed 2017-01-04] [WebCite Cache ID 6nGM8urHJ]
97. Norfolk Community Health and Care NHS Trust. Diabetes Type 2 education programme: Overview URL: http://www.norfolkcommunityhealthandcare.nhs.uk/The-care-we-offer/Service-search/diabetes-type-2-education-programme.htm [accessed 2016-07-27] [WebCite Cache ID 6jJhZ1LeZ]
98. Buckinghamshire Healthcare NHS Trust. About Us: National recognition for diabetes education service (Life and health with diabetes) 2013 URL: http://www.buckshealthcare.nhs.uk/About/national-recognition-for-diabetes-education-services.htm [accessed 2016-07-27] [WebCite Cache ID 6jJhuMltI]
99. Buckinghamshire Healthcare NHS Trust. Buckinghamshire Formulary: Management of type 2 diabetes URL: http://www.bucksformulary.nhs.uk/docs/Guideline_667FM.pdf [accessed 2017-01-04] [WebCite Cache ID 6nGML1mk2]
100. Bristol Community Health. Living with diabetes URL: http://briscomhealth.org.uk/diabetes-education-courses/ [accessed 2017-01-04] [WebCite Cache ID 6nGMTDrpr]
101. Newham Clinical Commissioning Group. Newham Community Prescription - what's it all about? URL: http://www.newhamccg.nhs.uk/GetInvolved/newham-community-prescription.htm [accessed 2016-07-27] [WebCite Cache ID 6jJne1kBA]
102. Newham Community Prescription URL: http://cohnewham.org/projects/newham-community-prescription/ [accessed 2017-01-04] [WebCite Cache ID 6nGMxyhmt]
103. Herbert N, Pilkington K, Ridge D, Dixon M, King D, Polley M. Feasibility of using facilitated social prescription to improve the health outcomes for those with Type 2 Diabetes and those at risk of diabetes. Eur J Integr Med 2015;7(6):688. [doi: 10.1016/j.eujim.2015.07.024]
104. College Surgery. Health facilitator URL: http://www.collegesurgery.org.uk/health_facilitator_p6619.html?a=0 [accessed 2016-07-27] [WebCite Cache ID 6jJpvBrVd]
105. Cambridge City Council. Fitness for adults with medical conditions (Start-Up exercise referral scheme) URL: https://www.cambridge.gov.uk/fitness-for-adults-with-medical-conditions [accessed 2017-01-04] [WebCite Cache ID 6nGN9gwq3]
106. Cambridge City Council. Cambridge City Council's Exercise Referral Schemes Annual Report 2014-15 URL: https://www.cambridge.gov.uk/sites/default/files/cambridge-exercise-referral-report-14-15-v2a.pdf [accessed 2017-01-04] [WebCite Cache ID 6nGNFrtIs]
107. Whitelaw S, Thirlwall C, Morrison A, Osborne J, Tattum L, Walker S. Developing and implementing a social prescribing initiative in primary care: insights into the possibility of normalisation and sustainability from a UK case study. Prim Health Care Res Dev 2016 Jun 8:1-10. [doi: 10.1017/S1463423616000219] [Medline: 27269818]

108. Choudhury S, Brophy S, Fareedi MA, Zaman B, Ahmed P, Williams DR. Intervention, recruitment and evaluation challenges in the Bangladeshi community: experience from a peer lead educational course. BMC Med Res Methodol 2008;8:64 [FREE Full text] [doi: 10.1186/1471-2288-8-64] [Medline: 18844992]
109. Choudhury SM, Brophy S, Fareedi MA, Zaman B, Ahmed P, Williams R. Examining the effectiveness of a peer-led education programme for Type 2 diabetes and cardiovascular disease in a Bangladeshi population. Diabet Med 2009 Jan;26(1):40-44. [doi: 10.1111/j.1464-5491.2008.02639.x] [Medline: 19125759]
110. NHS England. NHS Diabetes Prevention Programme (NHS DPP) URL: https://www.england.nhs.uk/ourwork/qual-clin-lead/diabetes-prevention/ [accessed 2016-12-02] [WebCite Cache ID 6mSCWr533]
111. Hiss RG. The concept of diabetes translation: addressing barriers to widespread adoption of new science into clinical care. Diabetes Care 2001 Jul;24(7):1293-1296. [Medline: 11423518]
112. Yates T, Davies M, Khunti K. Preventing type 2 diabetes: can we make the evidence work? Postgrad Med J 2009 Sep;85(1007):475-480. [doi: 10.1136/pgmj.2008.076166] [Medline: 19734515]
113. Yates T, Davies MJ, Henson J, Troughton J, Edwardson C, Gray LJ, et al. Walking away from type 2 diabetes: trial protocol of a cluster randomised controlled trial evaluating a structured education programme in those at high risk of developing type 2 diabetes. BMC Fam Pract 2012;13:46 [FREE Full text] [doi: 10.1186/1471-2296-13-46] [Medline: 22642610]
114. Linmans JJ, Spigt MG, Deneer L, Lucas AE, de Bakker M, Gidding LG, et al. Effect of lifestyle intervention for people with diabetes or prediabetes in real-world primary care: propensity score analysis. BMC Fam Pract 2011;12:95 [FREE Full text] [doi: 10.1186/1471-2296-12-95] [Medline: 21914190]
115. Dunkley AJ, Bodicoat DH, Greaves CJ, Russell C, Yates T, Davies MJ, et al. Diabetes prevention in the real world: effectiveness of pragmatic lifestyle interventions for the prevention of type 2 diabetes and of the impact of adherence to guideline recommendations: a systematic review and meta-analysis. Diabetes Care 2014 Apr;37(4):922-933. [doi: 10.2337/dc13-2195] [Medline: 24652723]
116. Glasgow RE, Emmons KM. How can we increase translation of research into practice? Types of evidence needed. Annu Rev Public Health 2007;28:413-433. [doi: 10.1146/annurev.publhealth.28.021406.144145] [Medline: 17150029]
117. Year of Care. Thanks for the Petunias: a guide to developing and commissioning non-traditional providers to support the self management of people with long term conditions 2011 URL: http://www.yearofcare.co.uk/sites/default/files/pdfs/Thanks%20for%20the%20Petunias.pdf [accessed 2016-11-07] [WebCite Cache ID 6lpGdylGs]
118. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Republished research: implementation research: what it is and how to do it: implementation research is a growing but not well understood field of health research that can contribute to more effective public health and clinical policies and programmes. This article provides a broad definition of implementation research and outlines key principles for how to do it. Br J Sports Med 2014 Apr;48(8):731-736. [doi: 10.1136/bmj.f6753] [Medline: 24659611]
119. Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney S, SQUIRE Development Group. Publication guidelines for quality improvement in health care: evolution of the SQUIRE project. Qual Saf Health Care 2008 Oct;17(Suppl 1):i3-i9 [FREE Full text] [doi: 10.1136/qshc.2008.029066] [Medline: 18836063]
120. Ogilvie D, Craig P, Griffin S, Macintyre S, Wareham NJ. A translational framework for public health research. BMC Public Health 2009;9:116 [FREE Full text] [doi: 10.1186/1471-2458-9-116] [Medline: 19400941]
121. 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]
122. Walach H, Falkenberg T, Fønnebø V, Lewith G, Jonas WB. Circular instead of hierarchical: methodological principles for the evaluation of complex interventions. BMC Med Res Methodol 2006;6:29 [FREE Full text] [doi: 10.1186/1471-2288-6-29] [Medline: 16796762]
123. Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. J Epidemiol Community Health 2002 Feb;56(2):119-127 [FREE Full text] [Medline: 11812811]
124. Husk K, Blockley K, Lovell R, Bethel A, Bloomfield D, Warber S, et al. What approaches to social prescribing work, for whom, and in what circumstances? A protocol for a realist review. Syst Rev 2016;5(1):93 [FREE Full text] [doi: 10.1186/s13643-016-0269-6] [Medline: 27255362]

Abbreviations

BMI: body mass index
BP: blood pressure
CHD: coronary heart disease
GP: general practitioner
HbA1c: Hemoglobin A1c
HCP: health care practitioners
LTC: long-term conditions
PP: published journal paper

RCT: randomized controlled trial
RR: risk ratio
(S)WEMWBS: (Short) Warwick Edinburgh Mental Wellbeing Scale
T2DM: type 2 diabetes
WR: Web-based report

Edited by G Eysenbach; submitted 02.08.16; peer-reviewed by R Simmons, MC Portillo; comments to author 08.09.16; revised version received 07.11.16; accepted 23.11.16; published 02.02.17.

Please cite as:
Pilkington K, Loef M, Polley M
Searching for Real-World Effectiveness of Health Care Innovations: Scoping Study of Social Prescribing for Diabetes
J Med Internet Res 2017;19(2):e20
URL: http://www.jmir.org/2017/2/e20/
doi:10.2196/jmir.6431
PMID:28153817

Original Paper

Patient Use of Electronic Prescription Refill and Secure Messaging and Its Association With Undetectable HIV Viral Load: A Retrospective Cohort Study

D Keith McInnes[1,2], MSc, ScD; Stephanie L Shimada[1,2,3], PhD; Amanda M Midboe[4], PhD; Kim M Nazi[5], PhD; Shibei Zhao[1], MPH; Justina Wu[4], MPH; Casey M Garvey[6], RN, BSN; Thomas K Houston[1,3], MD, MPH

[1]Center for Healthcare Organization and Implementation Research, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA, United States
[2]Department of Health Law, Policy, and Management, Boston University School of Public Health, Boston, MA, United States
[3]Division of Health Informatics and Implementation Science, Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States
[4]Center for Innovation to Implementation, VA Palo Alto Health Care System, Palo Alto, CA, United States
[5]Veterans and Consumers Health Informatics Office, Office of Connected Care, Veterans Health Administration, Washington, DC, United States
[6]School of Nursing, Bouve College of Health Sciences, Northeastern University, Boston, MA, United States

Corresponding Author:
D Keith McInnes, MSc, ScD
Center for Healthcare Organization and Implementation Research
Edith Nourse Rogers Memorial Veterans Hospital
Bldg 70
200 Springs Road
Bedford, MA, 01730
United States
Phone: 1 781 687 3507
Fax: 1 781 687 3106
Email: keith.mcinnes@va.gov

Abstract

Background: Electronic personal health records (PHRs) can support patient self-management of chronic conditions. Managing human immunodeficiency virus (HIV) viral load, through taking antiretroviral therapy (ART) is crucial to long term survival of persons with HIV. Many persons with HIV have difficulty adhering to their ART over long periods of time. PHRs contribute to chronic disease self-care and may help persons with HIV remain adherent to ART. Proportionally veterans with HIV are among the most active users of the US Department of Veterans Affairs (VA) PHR, called My HealtheVet. Little is known about whether the use of the PHR is associated with improved HIV outcomes in this population.

Objective: The objective of this study was to investigate whether there are associations between the use of PHR tools (electronic prescription refill and secure messaging [SM] with providers) and HIV viral load in US veterans.

Methods: We conducted a retrospective cohort study using data from the VA's electronic health record (EHR) and the PHR. We identified veterans in VA care from 2009-2012 who had HIV and who used the PHR. We examined which ones had achieved the positive outcome of suppressed HIV viral load, and whether achievement of this outcome was associated with electronic prescription refill or SM. From 18,913 veterans with HIV, there were 3374 who both had a detectable viral load in 2009 and who had had a follow-up viral load test in 2012. To assess relationships between electronic prescription refill and viral control, and SM and viral control, we fit a series of multivariable generalized estimating equation models, accounting for clustering in VA facilities. We adjusted for patient demographic and clinical characteristics associated with portal use. In the initial models, the predictor variables were included in dichotomous format. Subsequently, to evaluate a potential dose-effect, the predictor variables were included as ordinal variables.

Results: Among our sample of 3374 veterans with HIV who received VA care from 2009-2012, those who had transitioned from detectable HIV viral load in 2009 to undetectable viral load in 2012 tended to be older (*P*=.004), more likely to be white (*P*<.001), and less likely to have a substance use disorder, problem alcohol use, or psychosis (*P*=.006, *P*=.03, *P*=.004, respectively). There was a statistically significant positive association between use of electronic prescription refill and change in HIV viral load status from 2009-2012, from detectable to undetectable (OR 1.36, CI 1.11-1.66). There was a similar association between SM use and viral load status, but without achieving statistical significance (OR 1.28, CI 0.89-1.85). Analyses did not demonstrate a dose-response of prescription refill or SM use for change in viral load.

Conclusions: PHR use, specifically use of electronic prescription refill, was associated with greater control of HIV. Additional studies are needed to understand the mechanisms by which this may be occurring.

(J Med Internet Res 2017;19(2):e34) doi:10.2196/jmir.6932

KEYWORDS

health records, personal; HIV; viral load; electronic prescribing; electronic mail; secure messaging; self care; veterans

Introduction

Electronic health records (EHRs) are increasingly being adopted by hospitals, health plans, and other health care providers to improve the efficiency and effectiveness of health care delivery, meet provisions of the Affordable Care Act, and to qualify for Meaningful Use financial incentives [1-3]. Many of these EHRs include personal health records (PHRs, also known as “patient portals”), which enable patients to view parts of their medical record such as laboratory results, past and future appointments, and upcoming preventive care. Patients can also use PHRs to manage their own care and communicate with their health care providers including sending secure electronic messages to their clinical providers. A common PHR feature, electronic refill of prescriptions (hereafter called “Rx refill”), for example, may contribute to better medication adherence by ensuring that patients consistently have their medications on hand [4]. Similarly, there is evidence that secure messaging (SM) may contribute to increased likelihood of achieving chronic illness control, such as diabetes and hypertension, because of improved patient-provider communication [5-7]. These PHR self-management tools may be especially beneficial for persons with human immunodeficiency virus (HIV) who must carefully adhere to their combination antiretroviral therapy (ART), but also often have other chronic conditions that they must manage simultaneously [8].

Controlling the amount of HIV virus in the bloodstream is crucial to persons with HIV, and combination ART is highly effective at reducing HIV viral load when taken as prescribed. But the management of HIV is complex. Patients must carefully adhere to their ART regimen both to control viremia and to reduce the likelihood of drug resistance. Lab work is needed regularly to monitor HIV virus levels and the status of the immune system through CD4 cell counts [9]. Additionally, many persons with HIV have comorbid chronic mental and physical health conditions that may lead to drug-drug interactions or disease-drug interactions with ART [8]. This suggests that patients with HIV would benefit from frequent provider-patient communication and the ability to easily manage and order their medications [10-12]. PHR self-management tools, in particular Web-based prescription refills and SM, may facilitate patients’ adherence to ART and thus contribute to subsequent undetectable viral load [13,14]. However, to our knowledge no studies have examined whether the use of specific PHR features of prescription refill and SM are associated with improved clinical outcomes for persons with HIV.

This study examines use of the My HealtheVet PHR Rx refill and SM features by persons with HIV in a large integrated health care system. We sought to examine whether use of these tools was associated with undetectable viral load. To this end, we identified a cohort of HIV-infected veterans receiving US Department of Veterans Affairs (VA) health care and examined their patterns of My HealtheVet PHR use along with their laboratory results for HIV viral load status in 2 time periods.

Methods

Study Design

We conducted a retrospective cohort study, identifying veterans with HIV who used the My HealtheVet PHR and followed them to assess outcomes of suppression of viral load. The study was approved by the Bedford Massachusetts VA Medical Center Institutional Review Board.

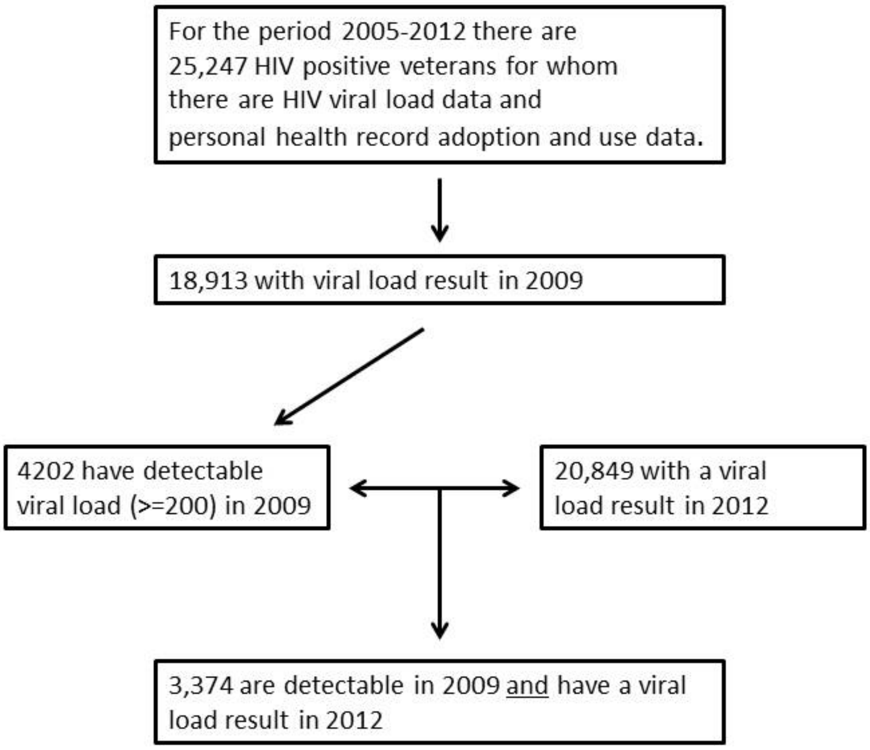
Setting

We used data from the VA system of medical records available through the VA Corporate Data Warehouse (CDW). Variables included patient demographics and International Classification of Disease, 9th revision, clinical modification (ICD-9-CM) diagnosis codes associated with all VA inpatient and outpatient encounters from October 1, 2007, to March 31, 2012. These data were linked at the patient level with My HealtheVet registration, SM, and Rx refill data from April 2012.

Sample

The study population included all American veterans aged 18 years and older who had obtained care from the VA health care system between April 1, 2010, and March 31, 2012 (N=6,012,875). Obtaining care in the VA was defined as having at least two outpatient visits or 1 inpatient hospitalization for any cause during this period. The cohort of HIV-infected patients was identified by examining VA’s decision support system laboratory results contained in CDW. Inclusion criteria for the cohort were: (1) a veteran determined to be HIV positive based on 2 or more instances of ICD-9-CM codes for HIV in the CDW, (2) had a detectable viral load in 2009, and (3) had viral load test results (detectable or undetectable) in 2012. For these analyses, an undetectable viral load was considered less than 200 copies of the HIV virus per milliliter (mL) of blood (<200/mL); conversely, 200 copies or more of the HIV virus per mL of blood (>=200/mL) indicated a detectable viral load. We identified 18,913 veterans who were HIV-positive and had a viral load test result in 2009 (See Figure 1). Of these, 4202 had a detectable viral load. Among these 4202 there were 3374 veterans who had a viral load test result in 2012. These 3374 veterans were the subject of this study. However, in the multivariable analyses, the total sample was 3289 due to missing values for 85 patients. For patients with multiple viral load tests in a given year, the most recent result in that calendar year (closest to December 31) was used to categorize the patient's viral load status.

Figure 1. Creation of analytical database.



VA's Personal Health Record, My HealtheVet

VA launched the My HealtheVet PHR in 2003 and enrollment has grown to approximately 3.5 million current registered users as of January 2016 [15]. The PHR provides evidence-based health information, health logs (eg, to track diet, exercise, weight), access to providers' notes, reminders for preventive health services (eg, immunization, cancer screening), prescription drug refills, and secure email messaging [16]. Since 2005, over 70 million prescription refills have been requested on the Web, and 1.7 million veterans have signed up for the SM service. Interestingly, veterans with HIV register for the My HealtheVet PHR at higher rates than veterans with other chronic illnesses. While 18.64% of all veterans receiving treatment from VA were registered to use the PHR in 2012, 26.48% of veterans with HIV were registered at the time [15].

Data Collection

Independent Variables

Use of My HealtheVet Rx refill and SM features were the independent variables of interest. We examined whether each of these tools was used by each veteran in our sample in the 3-year follow-up period from 2010-2012, and whether use of the tool was associated with undetectable viral load. These 2 variables were dichotomous with 1 indicating use of Rx refill at least one time in the 2010-12 period, and 0 indicating no use of Rx refill during that time period. The SM variable was created in the same way, with 1 indicating at least one use in 2010-12 and 0 indicating no use during that time period. In separate analyses, we also explored whether there may be a dose-response such that sustained use of the tool (Rx refill or SM) was associated with an increased likelihood of having an undetectable viral load. For those analyses, each of the variables (Rx refill and SM) was coded as an ordinal variable with 0 indicating no use in 2010-12, 1 indicating use in only 1 of the 3 years, and 2 indicating use in 2 or more of the 3 years.

Dependent Variable

The outcome of interest was viral load status in 2012. It was coded 1 for viral load < 200 mL (undetectable) and 0 for viral load >=200 mL (detectable).

Covariates

Following Shimada et al who examined electronic prescription refill and SM among 6 million veterans using VA health services, we included in our multivariable analyses the following demographic characteristics: age, gender, race or ethnicity, urban or rural residence based on home postal code, and economic need defined as eligibility for free care based on an annual VA means testing [15]. We grouped participants by age (<45 years or 45-54 years or 55-64 years or 65+ years), marital status (married vs never married or divorced or separated or widowed), and race or ethnicity (African American vs White or Asian or Native Hawaiian or American Indian or other). Chronic conditions and disease burden were measured using ICD-9-CM medical and mental health diagnoses that appeared at least once for an inpatient stay or at least twice for outpatient visits. The diagnostic grouping to identify chronic conditions had been previously validated in a large VA longitudinal study of veterans with HIV [17]. Comorbidities may have influenced both disease self-management behaviors and interest in using the PHR. We included comorbidities in our models in 2 ways. First, problem alcohol use, other substance use, depression, and psychoses were each included as independent variables, following their identification using ICD-9-CM codes, as described earlier. These 4 conditions were singled out because of their relatively high prevalence among HIV-infected populations [18,19] and their known negative association with medication adherence [20-22]. In addition, as an overall measure of chronic disease burden, we used the Elixhauser score [23-25]. This is a summary measure of disease burden that combines 31 common comorbidities. Because all patients in our sample were HIV positive, we adjusted the Elixhauser measure by removing HIV disease. In addition, we removed problem alcohol use, substance use, depression, and psychoses from the Elixhauser measure because they had already been included as independent variables as described previously.

Analyses

We began by examining bivariate relationships between HIV viral control and the independent variables using cross-tabulation and chi-square tests. To assess relationships between Rx refill and viral control, and SM and viral control, with the sample of 3289 patients, we fit a series of multivariable generalized estimating equation models, accounting for clustering in VA facilities. To understand how patient characteristics might bias the primary association of Rx refill (or SM) and viral control, our models controlled for socio-demographics (age, gender, race or ethnicity, economic need, marital status, rural or urban status) and health status (ie, Elixhauser comorbidity burden, as well as separate indicators of alcohol use, other substance use, depression, and psychoses). We ran models for Rx refill and SM separately. In initial models the predictor variables were included in dichotomous format. Subsequently, to evaluate a potential dose-effect, the predictor variables were included as ordinal, as described previously. This produced 4 different models. We considered statistical significance to be $P<.05$. To assess potential multicollinearity, we estimated variance inflation factor (VIF) for each independent variable in each of our multivariable models. No variables had VIF greater than 10 (the largest VIF was 1.61) suggesting there was little or no multicollinearity. This also suggests sampling bias was not a substantial issue. To assess whether being registered to use the PHR or not would affect our findings, we conducted sensitivity analyses in which we restricted the sample to the 1130 registered patients. When we ran the same 4 regression models, our results did not differ substantially from the original 4 models. Thus in the results section discussed later on, we focus on the analyses using the larger sample of 3289 patients.

Results

Patient Characteristics and PHR Use

Sociodemographic, health, and PHR use data are shown in Table 1, by viral load detectable status in 2012. Of the 3374 who had detectable viral loads in 2009, 66.60% had undetectable viral loads in 2012. Nearly three-quarters (71.61%) of the sample were in the 45-64 year age range, and 96.21% were male (consistent with the demographics of the US veteran population). The majority of the sample was African American (61.53%). Close to half (46.06%) were considered having high economic need, based on eligibility for free VA health care services due to low income, and 11.47% were married. Most of the sample (88.97%) lived in urban areas. There were considerable mental health and substance use disorder morbidity, including 60.04% with a diagnosis of depression, 41.44% with a substance use disorder, 32.99% with problem alcohol use, and 6.32% with psychoses. Excluding the aforementioned 4 diagnoses and HIV, subjects had a mean of 2.7 chronic conditions (as assessed with the Elixhauser score). One-third (33.49%) had registered for the PHR as of 2012. In addition 17.81% of the sample had used the PHR’s electronic prescription refill at least once in the 2010-2012 period, and 5.93% had sent at least one secure message to a VA provider during that period.

Table 1. Participant characteristics, overall and by human immunodeficiency virus (HIV) viral control in 2012.

Variable		Overall, n (%)	Viral control		*P* value
			Viral load undetectable, n (%)	Viral load detectable, n (%)	
		3374 (100)	2247 (66.60)	1127 (33.40)	
Age in years					.003
	<45	637 (18.88)	414 (18.42)	223 (19.79)	
	45-54	1190 (35.27)	776 (34.53)	414 (36.73)	
	55-64	1226 (36.34)	814 (36.23)	412 (36.56)	
	65+	321 (9.51)	243 (10.81)	78 (6.92)	
Gender					.05
	Female	128 (3.79)	75 (3.34)	53 (4.70)	
	Male	3246 (96.21)	2172 (96.66)	1074 (95.30)	
Race					<.001
	White	1130 (33.49)	826 (36.76)	304 (26.97)	
	Black	2076 (61.53)	1309 (58.26)	767 (68.06)	
	Hispanic	28 (0.83)	15 (0.67)	13 (1.15)	
	Native Hawaiian	29 (0.86)	24 (1.07)	5 (0.44)	
	American Indian	14 (0.41)	10 (0.45)	4 (0.35)	
	Asian	11 (0.33)	8 (0.36)	3 (0.27)	
	Unknown	86 (2.55)	55 (2.45)	31 (2.75)	
Economic need (means test)					.22
	Eligible for free care	1554 (46.06)	1018 (45.30)	536 (47.56)	
	Not eligible	1820 (53.94)	1229 (54.70)	591 (52.44)	
Marital status					.001
	Married	387 (11.47)	259 (11.53)	128 (11.36)	
	Never married	1567 (46.44)	1040 (46.28)	527 (46.76)	
	Divorced	1037 (30.74)	700 (31.15)	337 (29.90)	
	Separated	228 (6.76)	154 (6.85)	74 (6.57)	
	Widowed	120 (3.56)	83 (3.69)	37 (3.28)	
	Others	35 (1.04)	11 (0.49)	24 (2.13)	
Urban or rural location					.18
	Missing	82 (2.43)	51 (2.27)	31 (2.75)	
	Urban	3002 (88.97)	1990 (88.56)	1012 (89.80)	
	Rural	290 (8.60)	206 (9.17)	84 (7.45)	
Comorbidities					
	Depression (% yes)	2024 (60.04)	1325 (58.97)	699 (62.19)	.07
	Substance use (% yes)	1397 (41.44)	894 (39.79)	503 (44.75)	.006
	Problem alcohol use (% yes)	1112 (32.99)	713 (31.73)	399 (35.50)	.03
	Psychoses (% yes)	213 (6.32)	123 (5.47)	90 (8.01)	.004
	Elixhauser comorbidity score, mean (STD)	2.69 (2.35)	2.67 (2.38)	2.72 (2.31)	.57
My HealtheVet PHR Use					
	Registered to use My HealtheVet as of 2012	1130 (33.49)	785 (34.94)	345 (30.61)	.01
	Use Rx refill (2010-2012)	601 (17.81)	435 (19.36))	166 (14.73)	<.001
	Use secure messaging (2010-2012)	200 (5.93)	147 (6.54)	53 (4.70)	.03

There were differences between veterans with undetectable versus detectable viral load. Veterans with undetectable viral load were older (10.81% vs 6.92%, 65+ years), more likely to be male (96.66% vs 95.30%), white (36.76% vs 26.97%), and divorced (31.15% vs 29.90%). They were less likely than veterans with detectable viral load to have symptoms of depression (58.97% vs 62.19%), substance use disorder (39.79% vs 44.75%), problem alcohol use (31.73% vs 35.50%), or psychoses (5.47% vs 8.01%). Those with undetectable viral load were also more likely to be registered for My HealtheVet (34.94% vs 30.61%), to have used Rx refill (19.36% vs 14.73%), and to have used SM (6.54% vs 4.70%).

Multivariable Models

In our multivariable model examining use of Rx refill, there was a positive association between use of Rx refill and viral load control (Table 2). Veterans using Rx refill had 1.36 the odds (95% CI 1.11-1.66) of having an undetectable viral load compared with veterans who did not use Rx refill, after adjusting for sociodemographic and health variables. In addition, older age (Odds ratio, OR 1.01, 95% CI 1.00-1.02) and being white (OR 1.49, 95% CI 1.21-1.83) were independently and positively associated with undetectable viral load, whereas a diagnosis of psychosis (OR 0.66, 95% CI 0.46-0.95) was associated with reduced likelihood of undetectable viral load.

Table 2. Multivariable analysis of the odds of undetectable viral load in 2012 in relation to Rx refill use (n=3289).

Variable	Estimate	Odds ratio	95% CI	*P* value
Use Rx refill 2010-2012 (dichotomous)	0.3085	1.36	1.11-1.66	.003
Age	0.0132	1.01	1.00-1.02	.004
Male	0.2555	1.29	0.86-1.94	.22
White race	0.399	1.49	1.21-1.83	<.001
Economic need (means test)	-0.101	0.90	0.78-1.05	.19
Married	-0.053	0.95	0.78-1.16	.61
Elixhauser score	-0.009	0.99	0.96-1.03	.61
Problem alcohol use	-0.007	0.99	0.84-1.17	.93
Substance use	-0.059	0.94	0.76-1.18	.60
Depression	-0.067	0.94	0.79-1.11	.44
Psychoses	-0.417	0.66	0.46-0.95	.03
Rural	0.1506	1.16	0.86-1.57	.33

There were similar associations when viral load was modeled with SM as the predictor variable. However, SM did not achieve a statistically significant association with viral load status (Table 3). Veteran using SM had 1.28 the odds (95% CI 0.89-1.85) of having undetectable viral load compared with veterans who did not use SM, after adjusting for sociodemographic and health variables. Similar to the Rx refill model, older age (OR 1.01, 95% CI 1.00-1.02) and being white (OR 1.53, 95% CI 1.25-1.86) were independently and positively associated with undetectable viral load. Conversely, a diagnosis of psychosis (OR 0.66, 95% CI 0.46-0.95) was associated with reduced likelihood of having an undetectable viral load.

Table 3. Multivariable analysis of the odds of undetectable human immunodeficiency virus (HIV) viral load in 2012 in relation to secure messaging (SM) use (n=3289).

Variable	Estimate	Odds ratio	95% CI	*P* value
Use secure messaging	0.2462	1.279	0.89-1.85	.19
Age	0.0114	1.012	1.00-1.02	.01
Male	0.2748	1.316	0.87-1.99	.19
White race	0.4231	1.527	1.25-1.86	<.001
Economic need (means test)	-0.107	0.898	0.78-1.04	.15
Married	-0.063	0.939	0.77-1.15	.53
Elixhauser score	-0.008	0.992	0.96-1.03	.64
Problem alcohol use	-0.017	0.983	0.84-1.16	.83
Substance use	-0.073	0.929	0.75-1.16	.51
Depression	-0.057	0.944	0.80-1.12	.51
Psychoses	-0.416	0.659	0.46-0.95	.03
Rural	0.1509	1.163	0.86-1.58	.33

There was no evidence of a dose effect for either Rx refill or SM when treated as ordinal variables (Tables 4 and 5). The overall association of Rx refill with undetectable viral load, however, was still evident. Compared with veterans with 0 uses of Rx refill, those with Rx refill use in 1 year had 1.38 the odds of undetectable viral load (95% CI 1.02-1.87) and those with Rx refill use in 2+ years had nearly the same odds ratio (OR 1.35, 95% CI 1.07-1.71) (Table 4).

Table 4. Multivariable analysis of the odds of undetectable human immunodeficiency virus (HIV) viral load in 2012 in relation to Rx refill use (Rx refill included as ordinal variable), for assessment of dose effect (n=3289).

Variable	Estimate	Odds ratio	95% CI	*P* value
Use Rx refill 0 years (ref)	-	-	-	-
Use Rx refill 1 year	0.3194	1.376	1.02-1.87	.04
Use Rx refill 2-3 years	0.3016	1.352	1.07-1.71	.01
Age	0.0132	1.013	1.00-1.02	.004
Male	0.2556	1.291	0.86-1.94	.22
White race	0.3995	1.491	1.21-1.83	<.001
Economic need (means test)	-0.101	0.904	0.78-1.05	.19
Married	-0.052	0.949	0.78-1.16	.61
Elixhauser score	-0.009	0.991	0.96-1.03	.61
Problem alcohol use	-0.007	0.993	0.84-1.17	.93
Substance use	-0.059	0.943	0.76-1.17	.60
Depression	-0.067	0.935	0.79-1.11	.44
Psychoses	-0.417	0.659	0.46-0.95	.03
Rural	0.1501	1.162	0.86-1.57	.33

The findings for SM also indicated no evidence of dose effect. The odds ratio for SM use in 1 year was 1.37 (95% CI 0.88-2.13) and for SM use in 2+ years was 1.02 (95% CI 0.51-2.01) though neither was statistically significant (Table 5).

Table 5. Multivariable analysis of the odds of undetectable human immunodeficiency virus (HIV) viral load in 2012 in relation to secure messaging (SM) use (as ordinal variable), for assessment of dose effect (n=3289).

Variable	Estimate	Odds ratio	95% CI	*P* value
Secure messaging 0 years (reference)	-	-	-	-
Secure messaging 1 year	0.3157	1.371	0.88-2.13	.16
Secure messaging 2-3 years	0.0149	1.015	0.51-2.01	.97
Age	0.0114	1.012	1.00-1.02	.01
Male	0.2759	1.318	0.87-1.99	.19
White race	0.4252	1.53	1.25-1.87	<.001
Economic need (means test)	-0.106	0.9	0.77-1.04	.16
Married	-0.064	0.938	0.77-1.14	.53
Elixhauser score	-0.008	0.992	0.96-1.03	.62
Problem alcohol use	-0.017	0.983	0.84-1.16	.84
Substance use	-0.073	0.929	0.75-1.16	.51
Depression	-0.056	0.945	0.80-1.12	.51
Psychoses	-0.421	0.657	0.45-0.95	.03
Rural	0.1501	1.162	0.86-1.58	.34

Our sensitivity analyses, in which the regression models were reestimated using the restricted sample of the 1130 registered patients, did not yield any substantively different results than those presented in Tables 2-5- (data not shown).

Discussion

Principal Findings

We found that among veterans with HIV, there was a positive association between My HealtheVet PHR use and undetectable viral load. Specifically, veterans who had detectable viral loads in 2009, and who used the Rx refill function between 2010 and 2012, had 1.36 times the likelihood of undetectable viral load in 2012 compared with veterans who did not use Rx refill. There was no evidence of dose effect of either of Rx refill or SM on the likelihood of undetectable viral load. No observational study can prove causality, but we had features that support a causal argument—we identified a longitudinal relationship, but not a clear dose-response. Thus it is possible that PHR use is a marker for some unmeasured covariates, such as engagement with the health care system.

Limitations

Due to the study limitations, we cannot rule out that there may be other explanations for our main finding of an association between use of Rx refill and undetectable viral load. The most salient limitation is that as an observational study there is potential for confounding by indication, in that patients who are already activated to improve their health may also be more likely to try new tools, such as PHRs. It is possible that more empowered and self-efficacious patients decide to use a PHR, whereas patients who are less motivated and more challenged with self-management tasks do not. Self-efficacy and empowerment may be driving forces behind achievement of undetectable viral load, and not actually PHR use alone. However, by limiting our analytic sample to the HIV positive veterans who had uncontrolled viral load in 2009, we sought to minimize variation in self-efficacy and empowerment. Another important variable, which was not available in our dataset and thus not included in our models, is stigma. Data indicate that Web-based tools are seen as particularly valuable, and used more often, among persons with stigmatized health conditions than among persons with nonstigmatized conditions [26]. Degree of perceived stigma may have been affecting the results found in our study.

Additionally, the linkage between Web-based prescription refill and undetectable viral load is presumably mediated by proper medication taking and medication adherence. Our data did not permit us to evaluate medication adherence, or whether the PHR prescription refills were specifically for ART as compared with medications for other comorbid conditions (eg, for diabetes, hypertension). Additional studies are needed that address these limitations, for example by randomizing patients to PHR use, and by assessing ART adherence to see whether the trajectory from PHR use to undetectable viral load occurs though the expected taking of ART. Our sample was all veterans and mostly male, therefore our results may not be generalizable to nonveterans and to women. Another way that generalizability may be limited is that the VA is unlike many smaller health care systems in that it has succeeded in registering a large number of its patients for PHR use, approximately 3.5 million users as of 2016 [15]. This number is likely only exceeded in the United States by the number of users in the Kaiser Permanente health care system [27]. Additionally the VA's PHR has more features than many other PHR systems (including the recent addition of OpenNotes—that is, patient access to clinician notes [28]), which may be one of the reasons for the relatively high patient enrollment rates. While these unique features of the VA's PHR reduce generalizability, alternatively the sheer number of users makes the VA an ideal environment for examining the impact of PHRs on health care delivery, processes, and outcomes, even for conditions with relatively low prevalence such as HIV.

Comparison With Prior Work

Overall evidence on the effect of PHR use on care processes and outcomes for a variety of health conditions show mixed results [29]. While Zhou et al found an association between patient SM use and improved intermediate outcomes for hypertension, cholesterol, and diabetes [5], a clinical trial conducted by Wagner et al randomly assigning subjects to PHR access or no access found no effects of PHR access on hypertension control, though subanalyses of frequent PHR users showed they achieved reductions in diastolic blood pressure, compared with the infrequent PHR users [30]. In a retrospective study of patients with diabetes at the Cleveland Clinic, Tenforde et al found in adjusted multivariable analyses that in a PHR user group compared with nonuser group there was more hemoglobin A1c (HbA1c) testing and lower HbA1c values (7.0% in the PHR group vs 7.3% in the non-PHR group; $P<.01$) [31]. A study of veterans who were using the My HealtheVet PHR found that use of Rx refill and SM were both associated with improved low density lipoprotein cholesterol levels. It also found sustained use (use each year over 3 or more years) of SM was associated with improved HbA1c levels among veterans with diabetes, whereas sustained use of Rx refill was associated with improved blood pressure control among veterans with hypertension [32]. In yet another diabetes study conducted with 54 patients at Vanderbilt University Medical Center, Wade-Vuturo et al found that use of the PHR's SM feature (more vs fewer messages sent) was moderately associated with lower HbA1c values in Spearman correlation analyses, ($\rho=-.29$, $P=.04$) [33].

In HIV, similarly, the evidence is mixed, with several studies suggesting that PHR use may assist with care processes and outcomes, but at least one study finding no association. A 6-site study by Shade et al [34] used a serial cross-sectional design in which each site selected a different health IT intervention to improve HIV care. One site introduced a continuity of care summary accessible to patients via a portal. At that site, based on data from 500 patients, there was an increased odds of undetectable viral load from baseline (prior to patient continuity of care summary) to follow up (OR 1.36, 95% CI 1.09-1.71) [34,35]. Patients were selected to be representative of the clinic population without regard for whether they used the continuity of care summary. A study by Crouch et al [36] found that PHR use was associated with improved HIV control. The team identified 80 HIV+ veterans using a nonprobability quota sample, selected to be split evenly between using and not using VA's My HealtheVet PHR. There was a statistically significant association between My HealtheVet PHR use and undetectable viral load, with 95% having undetectable viral load in the PHR group compared with 70% in the non-PHR group ($P=.046$) [36].

This study did not examine which components of the PHR patients used. McInnes et al reported that in a cross-sectional sample of 1871 veterans, multivariable analyses found that self-reported PHR use was associated with pharmacy refill ART adherence [4]. The findings from logistic regression analyses by Gordon et al, in contrast to the other studies, found in comparing 39 PHR users with 43 nonusers that there was no association of PHR use with either self-reported ART adherence or medical record viral load control [37]. What the evidence in HIV and in other health conditions suggests is that there is greater need for randomized studies. This becomes more difficult as PHRs are increasingly included with EHR packages that hospitals and health systems adopt; however, there are still patient groups that have had little or no PHR exposure.

Our findings also point to potential racial disparities in access to and use of HIV related care. In each of our multivariable analyses, white patients had greater odds of achieving viral control than other races (which were overwhelmingly black). This may be due to uncontrolled variables such as education level which may vary by race. Other research, including studies in the VA health care system, suggests various sources for racial disparities. Richardson et al recently examined racial differences in HIV care (and in comorbid care for HIV-infected patients) in VA and reported, “Despite the lack of insurance-related barriers to care in the equal-access VHA health care system, racial disparities in the care for veterans with HIV remain problematic and extend to comorbid conditions” [38]. Whereas the knowledge about the causes of these disparities is incomplete, Richardson et al suggest that possible contributors include patient attitudes and beliefs, provider attitudes and implicit biases, differences in patient-provider communication, and social-structural disadvantages by race [39-42].

The persistent finding, in each of our models, of a negative association between the presence of psychoses and uncontrolled viral load is noteworthy. This may indicate patients in this group have difficulty adhering to their HIV medications. Research into nonadherence (in general, not just for HIV) for persons with psychoses suggests that there are a number of potential barriers to adherence to medications. These include lack of social support, problems with therapeutic alliance, lack of daily routines, negative attitudes toward medications, and cognitive deficits [43].

More broadly, among persons with HIV, adoption and use of PHRs have been associated with a number of sociodemographic characteristics. A study of veterans with HIV found that PHR use is associated with younger age, less than excellent or very good health, white race, more education, lack of substance use disorder, and higher incomes [4]. Patient PHR use also has some advantages over face-to-face communication. For health-related questions that are perceived by patients as sensitive or that may cause embarrassment or perceived stigma, PHRs provide a sense of privacy and anonymity [26,44].

PHRs provide patients with greater access to their providers and health information, and also provide tools that allow patients to undertake health self-management tasks more efficiently than through in-person or phone contact with their health care team [45-47]. Patients report that accessing information in their PHR can enhance communication with providers, for example through improved preparation for in-person visits. PHRs and the information they contain also aid patients by improving knowledge of their health, serving as reminders for needed follow up, and creating a sense of accountability for one’s health that leads to more self-care [48]. Accordingly, PHRs can be expected to enable patients to increase self-management of chronic conditions, resulting in better control of those conditions and better health overall. However, it is unclear whether simply providing a PHR to patients is sufficient to achieve this, as indicated by the results of the study by Wagner et al described previously [30]. With the provision of PHRs, there may be a need for patient orientation, training, and continued promotion to increase both adoption and sustained use, and to encourage specific uses that are likely to have the most health-related benefits.

Conclusions

Our examination of PHR use adds important information to the existing body of work. Our data come from a highly stigmatized and vulnerable population of veterans with HIV, many of whom have high economic need, are racial or ethnic minorities, have a mental health or substance use disorder, and may lack social support (only 11% married). Interestingly, this population has been shown to use VA’s My HealtheVet PHR more than other chronic disease groups [15]; however, the cause of this phenomenon is not completely clear. That the use of a PHR electronic prescription refill feature was positively associated with undetectable viral load status in this vulnerable population is encouraging in that it may indicate an augmentation of access to medications and providers. The PHR tools of Rx refill and SM may afford patients a greater sense of freedom to perform functions—ordering refills and communicating with providers—unencumbered by the risk of social stigma associated with face-to-face or phone contacts [26]. Reducing such stigma may encourage greater patient interactions with providers and the health care system when issues arise, and improve self-efficacy in addressing health-related tasks and challenges. This in turn should help reduce the occurrence of patients running out of prescription medications and also facilitate swift resolution of serious side effects a patient may be experiencing from medications or drug-drug interactions.

Given our observational study design, however, further examination of the potential benefits of PHRs, and tools such as electronic prescription refill, are merited, especially if they can involve randomization. That a considerable proportion of this population is using PHR tools is encouraging—and yet about two-thirds do not. This suggests that continued efforts are needed to reach out to this population and to provide eHealth tools that are seen as easy to use and beneficial, regardless of the background or socio-economic and health status of the potential user.

Acknowledgments

We thank Dr Allen Gifford and Dr Matthew Goetz for their advice on the interpretation of the corporate data warehouse viral load laboratory results and Dr Cynthia Brandt for consulting on methods for identifying HIV diagnoses in VA datasets. Dr McInnes was supported by a Career Development Award (CDA 09-016) from VA Health Services Research and Development Service, and 3 VA Quality Enhancement Research Initiatives: eHealth QUERI (EHQ 10-190), HIV/Hepatitis QUERI (HIV 98-001), and Bridging the Care Continuum QUERI (15-284). The contents do not represent the views of the US Department of Veterans Affairs or the US Government.

Conflicts of Interest

None declared.

References

1. Ralston JD, Hereford J, Carrell D. Use and satisfaction of a patient Web portal with a shared medical record between patients and providers. AMIA Annu Symp Proc 2006:1070 [FREE Full text] [Medline: 17238689]
2. Kahn JS, Hilton JF, Van Nunnery T, Leasure S, Bryant KM, Hare CB, et al. Personal health records in a public hospital: experience at the HIV/AIDS clinic at San Francisco General Hospital. J Am Med Inform Assoc 2010 Feb;17(2):224-228 [FREE Full text] [doi: 10.1136/jamia.2009.000315] [Medline: 20190069]
3. Bitton A, Flier LA, Jha AK. Health information technology in the era of care delivery reform: to what end? J Am Med Assoc 2012 Jun 27;307(24):2593-2594. [doi: 10.1001/jama.2012.6663] [Medline: 22735426]
4. McInnes DK, Shimada SL, Rao SR, Quill A, Duggal M, Gifford AL, et al. Personal health record use and its association with antiretroviral adherence: survey and medical record data from 1871 US veterans infected with HIV. AIDS Behav 2013 Jan 20;17(9):3091-3100. [doi: 10.1007/s10461-012-0399-3]
5. Zhou YY, Kanter MH, Wang JJ, Garrido T. Improved quality at Kaiser Permanente through e-mail between physicians and patients. Health Aff 2010 Jul 06;29(7):1370-1375. [doi: 10.1377/hlthaff.2010.0048]
6. Chen C, Garrido T, Chock D, Okawa G, Liang L. The Kaiser Permanente electronic health record: transforming and streamlining modalities of care. Health Aff 2009 Mar 01;28(2):323-333. [doi: 10.1377/hlthaff.28.2.323]
7. Price M, Bellwood P, Kitson N, Davies I, Weber J, Lau F. Conditions potentially sensitive to a personal health record (PHR) intervention, a systematic review. BMC Med Inform Decis Mak 2015;15:32 [FREE Full text] [doi: 10.1186/s12911-015-0159-1] [Medline: 25927384]
8. Holtzman C, Armon C, Tedaldi E, Chmiel JS, Buchacz K, Wood K, et al. Polypharmacy and risk of antiretroviral drug interactions among the aging HIV-infected population. J Gen Intern Med 2013 Apr 20;28(10):1302-1310. [doi: 10.1007/s11606-013-2449-6]
9. Yehia BR, Long JA, Stearns CR, French B, Tebas P, Frank I. Impact of transitioning from HIV clinical trials to routine medical care on clinical outcomes and patient perceptions. AIDS Care 2012 Jun;24(6):769-777. [doi: 10.1080/09540121.2011.630368]
10. Holtzer CD, Roland M. The use of combination antiretroviral therapy in HIV-infected patients. Ann Pharmacother 1999 Feb;33(2):198-209. [Medline: 10084416]
11. Montaner JS, Reiss P, Cooper D, Vella S, Harris M, Conway B, et al. A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for HIV-infected patients: the INCAS Trial. Italy, The Netherlands, Canada and Australia Study. J Am Med Assoc 1998 Mar 25;279(12):930-937. [Medline: 9544767]
12. Bangsberg DR, Hecht FM, Charlebois ED, Zolopa AR, Holodniy M, Sheiner L, et al. Adherence to protease inhibitors, HIV-1 viral load, and development of drug resistance in an indigent population. AIDS 2000 Mar 10;14(4):357-366. [Medline: 10770537]
13. Sarkar U, Lyles CR, Parker MM, Allen J, Nguyen R, Moffet HH, et al. Use of the refill function through an online patient portal is associated with improved adherence to statins in an integrated health system. Med Care 2014 Mar;52(3):194-201 [FREE Full text] [doi: 10.1097/MLR.0000000000000069] [Medline: 24374412]
14. Ralston JD, Cook AJ, Anderson ML, Catz SL, Fishman PA, Carlson J, et al. Home blood pressure monitoring, secure electronic messaging and medication intensification for improving hypertension control: a mediation analysis. Appl Clin Inform 2014;5(1):232-248 [FREE Full text] [doi: 10.4338/ACI-2013-10-RA-0079] [Medline: 24734136]
15. Shimada SL, Brandt CA, Feng H, McInnes DK, Rao SR, Rothendler JA, et al. Personal health record reach in the Veterans Health Administration: a cross-sectional analysis. J Med Internet Res 2014 Dec 12;16(12):e272 [FREE Full text] [doi: 10.2196/jmir.3751] [Medline: 25498515]
16. Nazi KM, Hogan TP, Wagner TH, McInnes DK, Smith BM, Haggstrom D, et al. Embracing a health services research perspective on personal health records: lessons learned from the VA My HealtheVet system. Journal of general internal medicine 2010 Jul;25(1):62-67 [FREE Full text] [doi: 10.1007/s11606-009-1114-6] [Medline: 20077154]
17. Goulet JL, Fultz SL, Rimland D, Butt A, Gibert C, Rodriguez-Barradas M, et al. Aging and infectious diseases: do patterns of comorbidity vary by HIV status, age, and HIV severity? Clin Infect Dis 2007 Dec 15;45(12):1593-1601 [FREE Full text] [doi: 10.1086/523577] [Medline: 18190322]

18. Bing EG, Burnam MA, Longshore D, Fleishman JA, Sherbourne CD, London AS, et al. Psychiatric disorders and drug use among human immunodeficiency virus-infected adults in the United States. Arch Gen Psychiatry 2001 Aug;58(8):721-728. [Medline: 11483137]
19. Justice AC, McGinnis KA, Atkinson JH, Heaton RK, Young C, Sadek J, Veterans Aging Cohort 5-Site Study Project Team. Psychiatric and neurocognitive disorders among HIV-positive and negative veterans in care: Veterans Aging Cohort Five-Site Study. AIDS 2004 Jan 01;18(Suppl 1):S49-S59. [Medline: 15075498]
20. Cené CW, Dennison CR, Powell HW, Levine D, Bone LR, Hill MN. Antihypertensive medication nonadherence in black men: direct and mediating effects of depressive symptoms, psychosocial stressors, and substance use. J Clin Hypertens (Greenwich) 2013 Mar;15(3):201-209 [FREE Full text] [doi: 10.1111/jch.12056] [Medline: 23458593]
21. Malow R, Dévieux JG, Stein JA, Rosenberg R, Jean-Gilles M, Attonito J, et al. Depression, substance abuse and other contextual predictors of adherence to antiretroviral therapy (ART) among Haitians. AIDS Behav 2013 May;17(4):1221-1230. [doi: 10.1007/s10461-012-0400-1] [Medline: 23338563]
22. Tucker JS, Burnam MA, Sherbourne CD, Kung F, Gifford AL. Substance use and mental health correlates of nonadherence to antiretroviral medications in a sample of patients with human immunodeficiency virus infection. Am J Med 2003 May;114(7):573-580. [Medline: 12753881]
23. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care 1998 Jun 14;36(1):8-27. [doi: 10.1111/mec.13956] [Medline: 9431328]
24. Menendez ME, Neuhaus V, van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clin Orthop Relat Res 2014 Sep;472(9):2878-2886 [FREE Full text] [doi: 10.1007/s11999-014-3686-7] [Medline: 24867450]
25. Austin SR, Wong Y, Uzzo RG, Beck JR, Egleston BL. Why summary comorbidity measures such as the Charlson Comorbidity Index and Elixhauser Score work. Med Care 2015 Sep;53(9):e65-e72 [FREE Full text] [doi: 10.1097/MLR.0b013e318297429c] [Medline: 23703645]
26. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
27. Johnson LW, Garrido T, Christensen K, Handley M. Successful practices in the use of secure e-mail. Perm J 2014 Jul;18(3):50-54 [FREE Full text] [doi: 10.7812/TPP/13-160] [Medline: 24937149]
28. Nazi KM, Turvey CL, Klein DM, Hogan TP, Woods SS. VA OpenNotes: exploring the experiences of early patient adopters with access to clinical notes. J Am Med Inform Assoc 2014 Oct 28:380-389. [doi: 10.1136/amiajnl-2014-003144] [Medline: 25352570]
29. 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;159(10):677-687. [Medline: 25973514]
30. Wagner PJ, Dias J, Howard S, Kintziger KW, Hudson MF, Seol Y, et al. Personal health records and hypertension control: a randomized trial. J Am Med Inform Assoc 2012 Jul;19(4):626-634 [FREE Full text] [doi: 10.1136/amiajnl-2011-000349] [Medline: 22234404]
31. Tenforde M, Nowacki A, Jain A, Hickner J. The association between personal health record use and diabetes quality measures. J Gen Intern Med 2012 Apr;27(4):420-424 [FREE Full text] [doi: 10.1007/s11606-011-1889-0] [Medline: 22005937]
32. Shimada SL, Allison JJ, Rosen AK, Feng H, Houston TK. Sustained use of patient portal features and improvements in diabetes physiological measures. J Med Internet Res 2016 Jul 01;18(7):e179 [FREE Full text] [doi: 10.2196/jmir.5663] [Medline: 27369696]
33. 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]
34. Shade SB, Steward WT, Koester KA, Chakravarty D, Myers JJ. Health information technology interventions enhance care completion, engagement in HIV care and treatment, and viral suppression among HIV-infected patients in publicly funded settings. J Am Med Inform Assoc 2015 Apr;22(e1):e104-e111 [FREE Full text] [doi: 10.1136/amiajnl-2013-002623] [Medline: 25030033]
35. Shade SB, Chakravarty D, Koester KA, Steward WT, Myers JJ. Health information exchange interventions can enhance quality and continuity of HIV care. Int J Med Inform 2012 Oct;81(10):e1-e9. [doi: 10.1016/j.ijmedinf.2012.07.003] [Medline: 22854158]
36. Crouch PB, Rose CD, Johnson M, Janson SL. A pilot study to evaluate the magnitude of association of the use of electronic personal health records with patient activation and empowerment in HIV-infected veterans. PeerJ 2015;3:e852 [FREE Full text] [doi: 10.7717/peerj.852] [Medline: 25802815]
37. Gordon P, Camhi E, Hesse R, Odlum M, Schnall R, Rodriguez M, et al. Processes and outcomes of developing a continuity of care document for use as a personal health record by people living with HIV/AIDS in New York City. Int J Med Inform 2012 Oct;81(10):e63-e73 [FREE Full text] [doi: 10.1016/j.ijmedinf.2012.06.004] [Medline: 22841825]

38. Richardson KK, Bokhour B, McInnes DK, Yakovchenko V, Okwara L, Midboe AM, et al. Racial disparities in HIV care extend to common comorbidities: implications for implementation of interventions to reduce disparities in HIV care. J Natl Med Assoc 2016;108(4):201-210.e3. [doi: 10.1016/j.jnma.2016.08.001]
39. Saha S, Jacobs EA, Moore RD, Beach MC. Trust in physicians and racial disparities in HIV care. AIDS Patient Care and STDs 2010 Jul;24(7):415-420. [doi: 10.1089/apc.2009.0288]
40. Burgess DJ, Warren J, Phelan S, Dovidio J, van Ryn M. Stereotype threat and health disparities: what medical educators and future physicians need to know. J Gen Intern Med 2010 Mar 30;25(S2):169-177. [doi: 10.1007/s11606-009-1221-4]
41. Perry M, Williams RL, Wallerstein N, Waitzkin H. Social capital and health care experiences among low-income individuals. Am J Public Health 2008 Feb;98(2):330-336. [doi: 10.2105/AJPH.2006.086306]
42. McDoom MM, Bokhour B, Sullivan M, Drainoni M. How older black women perceive the effects of stigma and social support on engagement in HIV care. AIDS Patient Care and STDs 2015 Feb;29(2):95-101. [doi: 10.1089/apc.2014.0184]
43. Velligan DI, Weiden PJ, Sajatovic M, Scott J, Carpenter D, Ross R, et al. Assessment of adherence problems in patients with serious and persistent mental illness: recommendations from the Expert Consensus Guidelines. J Psychiatr Pract 2010 Jan;16(1):34-45. [doi: 10.1097/01.pra.0000367776.96012.ca] [Medline: 20098229]
44. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
45. Abdullah M, Theobald DE, Butler D, Kroenke K, Perkins A, Edgerton S, et al. Access to communication technologies in a sample of cancer patients: an urban and rural survey. BMC Cancer 2005 Feb 17;5:18 [FREE Full text] [doi: 10.1186/1471-2407-5-18] [Medline: 15717933]
46. Horsky J, Morgan SJ, Ramelson HZ. Coordination of care for complex pediatric patients: perspectives from providers and parents. AMIA Annu Symp Proc 2014 Nov 14:681-690.
47. 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]
48. 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 Mar 27;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]

Abbreviations

ART: antiretroviral therapy
CDW: Corporate Data Warehouse
EHR: electronic health record
HbA1C: Hemoglobin A1c
HIV: human immunodeficiency virus
IT: information technology
mL: milliliter
OR: odds ratio
PHR: personal health record
SM: secure messaging
VA: Department of Veterans Affairs

Edited by G Eysenbach; submitted 01.11.16; peer-reviewed by K Turner, C Turvey; comments to author 18.11.16; revised version received 30.12.16; accepted 03.01.17; published 15.02.17.

Please cite as:
McInnes DK, Shimada SL, Midboe AM, Nazi KM, Zhao S, Wu J, Garvey CM, Houston TK
Patient Use of Electronic Prescription Refill and Secure Messaging and Its Association With Undetectable HIV Viral Load: A Retrospective Cohort Study
J Med Internet Res 2017;19(2):e34
URL: http://www.jmir.org/2017/2/e34/
doi:10.2196/jmir.6932
PMID:28202428

Original Paper

Enhancing Doctors' Competencies in Communication With and Activation of Older Patients: The Promoting Active Aging (PRACTA) Computer-Based Intervention Study

Dorota Wlodarczyk[1], PhD; Joanna Chylińska[1], PhD; Magdalena Lazarewicz[1], PhD; Marta Rzadkiewicz[1], PhD; Mariusz Jaworski[1], PhD; Miroslawa Adamus[1], PhD; Gørill Haugan[2], PhD; Monica Lillefjell[3], PhD; Geir Arild Espnes[2], PhD

[1] Second Faculty of Medicine, Department of Medical Psychology, Medical University of Warsaw, Warsaw, Poland

[2] NTNU Center for Health Promotion Research, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway

[3] NTNU Center for Health Promotion Research, Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding Author:
Dorota Wlodarczyk, PhD
Second Faculty of Medicine
Department of Medical Psychology
Medical University of Warsaw
ZJAM Bldg, 1st Fl
81 Zwirki i Wigury St
Warsaw, 02-091
Poland
Phone: 48 501582248
Fax: 48 22 5720 542
Email: dorota@wde.com.pl

Abstract

Background: Demographic changes over the past decades call for the promotion of health and disease prevention for older patients, as well as strategies to enhance their independence, productivity, and quality of life.

Objective: Our objective was to examine the effects of a computer-based educational intervention designed for general practitioners (GPs) to promote active aging.

Methods: The Promoting Active Aging (PRACTA) study consisted of a baseline questionnaire, implementation of an intervention, and a follow-up questionnaire that was administered 1 month after the intervention. A total of 151 primary care facilities (response rate 151/767, 19.7%) and 503 GPs (response rate 503/996, 50.5%) agreed to participate in the baseline assessment. At the follow-up, 393 GPs filled in the questionnaires (response rate, 393/503, 78.1%), but not all of them took part in the intervention. The final study group of 225 GPs participated in 3 study conditions: e-learning (knowledge plus skills modelling, n=42), a pdf article (knowledge only, n=89), and control (no intervention, n=94). We measured the outcome as scores on the Patients Expectations Scale, Communication Scale, Attitude Toward Treatment and Health Scale, and Self-Efficacy Scale.

Results: GPs participating in e-learning demonstrated a significant rise in their perception of older patients' expectations for disease explanation (Wald χ^2=19.7, P<.001) and in perception of motivational aspect of older patients' attitude toward treatment and health (Wald χ^2=8.9, P=.03) in comparison with both the control and pdf article groups. We observed additional between-group differences at the level of statistical trend. GPs participating in the pdf article intervention demonstrated a decline in self-assessed communication, both at the level of global scoring (Wald χ^2=34.5, P<.001) and at the level of 20 of 26 specific behaviors (all P<.05). Factors moderating the effects of the intervention were the number of patients per GP and the facility's organizational structure.

Conclusions: Both methods were suitable, but in different areas and under different conditions. The key benefit of the pdf article intervention was raising doctors' reflection on limitations in their communication skills, whereas e-learning was more effective in changing their perception of older patients' proactive attitude, especially among GPs working in privately owned facilities and having a greater number of assigned patients. Although we did not achieve all expected effects of the PRACTA intervention, both its forms seem promising in terms of enhancing the competencies of doctors in communication with and activation of older patients.

(J Med Internet Res 2017;19(2):e45) doi:10.2196/jmir.6948

KEYWORDS

health services for the aged; active aging; e-learning; general practioners; professional competence; psychosocial competencies; health communication; seniors' expectations; seniors' attitude toward treatment and health

Introduction

Why Is Activation of Older Patients Important While Practicing Medicine?

Demographic changes over the past decades have created new challenges for health care providers [1]. Current trends in aging, along with continued limitations in providing effective health care to older patients, make it necessary to offer activities that promote health and disease prevention for older people, as well as strategies to enhance their independence, productivity, and quality of life [2,3].

Older patients' use of health care services is increasing [3], but certain aging processes may be modified by individual activity. Thus, a general practitioner's (GP) office appears to be the right place to discuss desired health- and activity-enhancing behaviors. Patients have been shown to depend on clinicians in areas such as preserving and promoting physical and emotional health [4]. Individuals who had been helped by a health provider to learn how to monitor their condition and to set goals were more activated than those who didn't receive such assistance [5,6]. At the same time, preventive health behavior and related matters have rarely been addressed by physicians and patients over 60 years of age in the primary health care setting [7].

Moreover, the lack of supporting patient activation and engagement was determined to be a potential pitfall for institutions aiming at improving quality and reducing health care costs [8].

What Might Activation of Older Patients by GPs Involve?

Patient activation is defined as the patient's ability to self-manage symptoms and problems, involvement in clinical decision making, and engagement in activities that maintain functioning and reduce health decline [9]. Its measurement allows their doctors to assign patients to 1 of 4 levels of activation, ranging from passive to proactive [10,11].

The unidimensional approach, however, does not allow for estimating the level of certain components of attitude toward treatment and health (ATH), such as cognitive, emotional (including positive and negative aspects), and motivational components [12,13]. An increasing number of studies have also shown the significance of self-efficacy in initiating and maintaining health behaviors [14].

To successfully activate older patients, it is necessary to identify their expectations of both the medical encounter [15] and effective communication [16,17]. Initial attempts to categorize patients' needs demonstrated the importance of explaining disease states, providing emotional support, and acting on or providing information about medical treatment [18]. As far as successful and active aging is concerned, the need for information about health promotion opportunities and improvement of quality of life is growing in importance [19].

In turn, accurate recognition of older patients' beliefs about their GPs' communication is one of the most important factors contributing to good physician-patient rapport [20,21]. Patient-centered communication, taking into account older people's perspective and empowering them in the process of decision making, was found to be associated with more favorable outcomes for patients [22,23]. Communicational aspects of the encounter are frequently rated by patients as the most important, although specificity of expectations depending on the medical situation has been found [24].

People over 65 years of age are seen through the prism of age-related stereotypes picturing them as poor, frail, and dependent [25-27]. Furthermore, there is an increasing number of studies on GPs' beliefs concerning older patients [28,29]. Such studies have taught us that physicians' attitudes toward elderly patients are diverse and positively related to factors such as geriatrics training, which suggests that providing GPs with appropriate knowledge and skills is essential [30]. Ageist misconceptions of older patients as indifferent, rigid in their beliefs and practices, or frequently annoyed may have serious negative consequences for them [31].

What Do We Know About E-Learning Versus Traditional Learning of GPs?

Selecting a teaching approach that both is relevant for the group and serves educational objectives is of significant importance, especially in the case of doctors, who are frequently overworked, lack time, and face changing demands. E-learning can be an attractive alternative to traditional learning. It is defined as any educational intervention that is mediated electronically via the Internet asynchronously [32].

The benefits of e-learning have been reported in terms of increased accessibility to education, improved self-efficacy, knowledge generation, cost effectiveness, learner flexibility, and interactivity; reports regarding its usage and effectiveness, however, have ranged broadly [33]. E-learning has proved to be at least as effective as traditional learning in terms of knowledge acquisition and user satisfaction, although knowledge about its effects on behavior change and patient outcomes is still insufficient [32]. A systematic review led to the conclusion that situated e-learning (embedded within a specific context representing real practice) enhances novice learners' performance, but that its effect on knowledge improvement is limited when compared with traditional learning [34]. There is also little knowledge about e-learning effectiveness in teaching GPs psychosocial competencies, such as communication and older patient activation. Moreover, little has been done to improve the knowledge about factors that moderate outcomes of different types of interventions [33].

Research Problems

Considering the above, we developed and investigated the Promoting Active Aging (PRACTA) intervention for GPs. It was designed to stimulate better recognition of older patients' expectations related to the medical encounter, more effective and patient-centered communication with older patients, and a better ability to enhance active ATH among older patients. We rooted the content of the intervention in selected theories in the field of health psychology, such as successful aging theories [35,36], attitude toward health [12], patient-centered communication [37], models of health behavior change [38], or social support theories [39], as well as the results of surveys from doctors and patients collected at time 1 of this study.

We presented the intervention in 2 forms: an e-learning platform and an article in pdf format. The e-learning intervention was a multimedia program aimed at presenting knowledge and modelling skills (demonstration of recommended solutions by means of, for example, case studies and video-recorded dialogues, and testing of new skills in simulated situations). The pdf article was a digitized text with pictures aimed solely at presenting knowledge. First, we anticipated that gaining some knowledge about older patients' expectations and ways of enhancing their active attitude toward health might contribute to a more positive perception of older patient activation among GPs, whereas practicing skills might cause more changes in GPs' behavior, thus initiating a positive spiral of changes (GPs use the skills of older patient activation, leading to older patients becoming more active, as a result of which GPs appraise older adults as being capable of being active, and consequently GPs continue older patient activation). Second, we assumed that, due to adoption of a wider range of didactic tools, e-learning would have a stronger and broader effect (in terms of number of outcome variables) than the pdf article.

We enrolled doctors in 1 of 3 study groups: e-learning, pdf article (comparative group), and control (no intervention at this time). The aim of the study was to examine the short-term effects of the PRACTA educational intervention in reference to the following outcome variables: (1) GPs' perception of older patients' expectations of the medical encounter, (2) GPs' self-assessed communication skills, and (3) GPs' perception of older patients' ATH.

We formulated 2 research questions (RQs), as follows. RQ1: Do the study groups differ from each other in terms of changes observed in the outcome variables? We hypothesized that, among the GPs in the intervention groups (e-learning and pdf article), in comparison with the control group, there would be (1a) a greater increase in their perception of older patients' expectations related to different types of information, with the increase in the e-learning group being larger than in the pdf article group (we had no directional hypotheses about other types of expectations), (1b) a greater increase in self-assessed communication skills, with the increase in the e-learning group being larger than in the pdf article group, (1c) a greater increase in their perception of older patients' active ATH, with the increase in the e-learning group being larger than in the pdf article group. RQ2: Are there any factors affecting the effects of the PRACTA intervention? Due to the exploratory nature of the question, we did not formulate any directional hypotheses but assumed that some sociodemographic, vocational, and organizational factors that we took into account would be moderators of the effects of the intervention.

Methods

Procedure and Recruitment

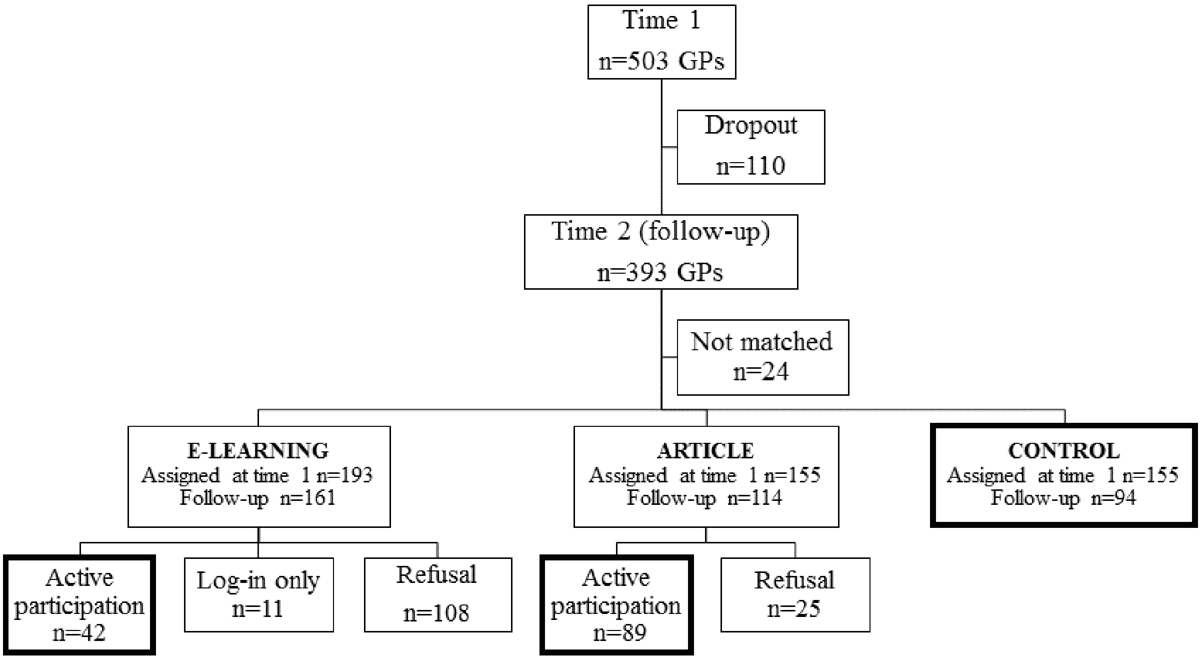
The PRACTA study consisted of a baseline questionnaire administered to GPs (time 1), implementation of the intervention (available for 3 months), and follow-up questionnaire administered to the GPs (time 2, which took place 1 month after the intervention). Recruitment comprised 2 stages: recruitment of primary care facilities (we contacted the facility's management to obtain permission to recruit the GPs) and recruitment of GPs working in these facilities. We considered the following inclusion criteria for facilities: delivering primary care, having a contract with the National Health Fund (patients did not pay for services out of their private funds), and being located in central Poland (a slightly wealthier part of the country including both urban and rural areas). The inclusion criteria for doctors were as follows: working in a facility recruited for the study, delivering primary care, and signing a written consent to participate in all parts of the project. The procedure guaranteed that data collection would be depersonalized, and every single GP was instructed in how to create an individual code that enabled matching scores from time 1 and time 2.

A total of 151 primary care facilities (response rate 151/767, 19.7%) and 503 GPs (response rate 503/996, 50.5%) agreed to participate in the baseline assessment. The facilities were randomly assigned to 3 groups: e-learning, pdf article, and control (we used random assignment of facilities only to ensure that all GPs working in the facility had access exclusively to one type of intervention). At time 2, there was a 78.1% response rate (393/503), but in 24 cases a missing or wrong individual code made matching scores from both measurements impossible (we considered these cases as dropouts). Figure 1 presents a flowchart of GP participation in consecutive stages of the project.

The final study sample consisted of 225 GPs: 42 actively taking part in e-learning (logged in and received points in at least one test), 89 actively participating in the pdf article intervention (filling out the form with questions regarding the pdf article that was an indicator of active participation), and 94 constituting the control group (participating in time 1 and time 2 surveys without any intervention at this time). As Figure 1 shows, 11 (21%) of 53 GPs logged in only, without receiving any points, but of the remaining 42 who obtained some points, 30 (71%) had completed all 5 modules, 2 (5%) had completed 4 modules, and 10 (24%) had completed 3 or fewer modules. The pdf article group was asked about what part of the article they had read: 47 (53%) declared that they had read the article in full, 19 (21%) had read three-quarters of the text, 10 (11%) had read approximately half of the article, and 13 (15%) had read approximately one-quarter or less. Whenever possible, interviewers asked doctors why they had declined to participate in the intervention, and “lack of time” was the most frequent response.

All GPs from the pdf article and control groups who declared their willingness to participate in PRACTA e-learning got access to it after completing the time 2 assessment. Doctors were interviewed by professional interviewers who had been trained on project-specific requirements for standardization of assessments. The content and form of both types of intervention were developed by researchers who prepared the whole project and carried out time 1 analyses. Approval for the study was obtained from the Bioethics Committee of the Medical University of Warsaw.

Figure 1. Flow chart of participation in the Promoting Active Aging (PRACTA) intervention study. GP: general practitioner.



PRACTA Intervention

We developed the PRACTA intervention in 2 forms. First, we prepared the e-learning intervention and then, based on its content, we created a pdf article. Both forms included 5 modules that were identically themed and presented in the same order; they were different, however, in their range, volume, and teaching approaches to present both knowledge and skills.

The modules covered the following subjects: (1) the process of active aging and the importance of an active attitude toward health, (2) doctors’ beliefs about older adults’ abilities and expectations, (3) the importance of physician-patient rapport for older patients and their health outcomes, (4) psychological rules and skills for promoting an active attitude toward health, and (5) quality of life and providing support for older patients (see Multimedia Appendix 1 for a detailed description of both types of intervention).

E-learning was designed to be a game in which participating players chose their character (female or male) and then receive specific task missions to complete. It included various multimedia components, which allowed for demonstration of specific practical solutions and for modelling communication and older adults’ activation skills. One module took about 1 hour to compete. To join the e-learning intervention, each participant was given a personal log-in and password and a USB flash drive with the information about time to access and rules of conduct.

The pdf article intervention took the form of a text with concise information, divided into small sections, and structured visually with simple pictures and figures (all images used in the pdf version were extracted from the e-learning intervention). Information presented in the pdf article had the form of a summary of the e-learning content and included a general description of solutions and techniques. Each pdf article module had a length of 3 pages of A4-sized paper. To join this form of intervention, each participant was given a USB flash drive with the article in pdf format.

Before the intervention started, we had launched an information campaign in primary care facilities (through letters sent to facility heads and their representatives who would be responsible for the PRACTA project-facility collaboration) and among GPs. The campaign for GPs comprised 2 phases: in phase 1, we mailed letters to GPs informing them about the timetable of scheduled stages of the project; in phase 2, interviewers visited GPs to deliver verbal information, leaflets, and USB flash drives with instructions and personal log-ins and passwords. After 4 weeks of the intervention (and evaluation of the response rate), we took other steps to remind and encourage GPs to participate: a letter, a visit by an interviewer, 2-3 telephone calls to a facility representative (until GPs clearly refused to participate) in the e-learning group and a letter only in the pdf article group.

The PRACTA e-learning intervention was administrated via the Medical University of Warsaw website. The e-learning platform (Microsoft SharePoint with SharePoint Learning Kit, version SLK 1.8; Microsoft Corporation) consisted of 2 systems: a learning content management system and a learning management system. To create the e-learning intervention, we used Articulate Storyline software (version 1; Articulate Global, Inc). To access PRACTA e-learning, each GP was given personal log-in and password (with information on time to access, rules of conduct, and availability of technical support). This procedure was very simple and user friendly (there were no requests for assistance). The first page after logging in contained overall information on the structure and format of the entire program. Participants were instructed to begin with module 1 and continue sequentially to finish with module 5. It was possible to stop and resume an uncompleted module at any time (the module was not available to a user only after completing the final test and saving scores). The participants were guided through the PRACTA e-learning by active arrows and written or audio instructions (eg, choose your character, click the button or read the materials, watch a video) with no interactions with a tutor or other participants (unidirectional). Navigating e-learning was easy and intuitive, allowing the users not to concentrate on technical aspects. The system allowed for monitoring the GPs' log-ins and test results evaluating their knowledge after completion of each module (only after saving the results).

The total number of points was calculated and converted into credit points (registered by the Polish Chamber of Physicians, *Izba Lekarska*, as a form of professional development). A diploma confirming the number of credit points was a major reward for participating in e-learning (certificates of participation in the PRACTA project and small project gadgets were given to all participants).

Measurement

The outcome variables in the study were (1) the GPs' perception of older patients' appointment-related expectations, (2) GPs' self-assessed communication behavior, and (3) the GPs' perception of older patients' ATH. The tools were designed for the purpose of the PRACTA project, and their psychometric properties were evaluated after the pilot study had been completed, and later, at the time of both questionnaires (time 1 and time 2) of the study proper. The questionnaires were originally written and administered in Polish, but their English translations are presented here in the multimedia appendixes.

GP's Perception of Older Patients' Appointment-Related Expectations

We measured this outcome variable with the PRACTA Patients Expectations Scale-Doctors (PRACTA-PES-D) consisting of six 3-item sections (each section was a single-factor scale) (Multimedia Appendix 2 shows the full scale in English translation). Each section reflected a different expectation: disease explanation, treatment explanation, health advice, quality-of-life improvement, rapport, and emotional support. Example items from each section are as follows: "Usually, during a visit elderly patients (65+) expect me…to find the cause of their symptoms" (disease explanation), "to present the results of the tests performed" (treatment explanation), "to encourage them to make health promoting changes" (health advice), "to suggest ways of maintaining life satisfaction" (quality-of-life improvement)," to show them respect" (rapport), and "to talk to them about how they feel and how they cope" (emotional support). The doctors assessed the importance of these expectations of older patients on a Likert scale from 1 to 7 (1=completely irrelevant, 7=very relevant). Each section score was calculated as a mean value (the sum of item scores divided by the number of items in a given section). The possible section score is between 1 and 7. The higher the score, the more significant the patient's expectation, as perceived by the GP. Reliability coefficients of PRACTA-PES-D scales were good or very good and ranged from .79 to .96 at time 1 and from .82 to .95 at time 2.

GPs' Self-Assessed Communication Behaviors

We assessed communication behavior with the PRACTA Communication Scale-Doctors (PRACTA-CS-D), comprising 26 items that allowed for calculating a global score or that could be analyzed separately (Multimedia Appendix 3 shows the full scale in English translation). GPs rated the frequency of communicational behavior presented in each item (eg, "During visits of my elderly patients (65+) I…greet them in a kind manner," "listen to them carefully," "make sure I understood them correctly,"). The answers were provided on a 7-point Likert scale, where 1=very seldom and 7=very often. The global score was calculated as a mean value of all item scores. It ranged from 1 to 7. The higher the score, the higher the frequency of communication behavior declared by GPs. Reliability coefficients of global scores before and after the intervention were .94 and .95, respectively.

GPs' Perception of Older Patients' ATH

We evaluated ATH by 2 tools: the PRACTA Attitude Toward Treatment and Health Scale-Doctors (PRACTA-ATH-D) and the PRACTA Self-Efficacy Scale-Doctors (PRACTA-SE-D) (Multimedia Appendix 4 to see all items measuring ATH: PRACTA-SE-D items directly follow PRACTA-ATH-D items). PRACTA-ATH-D comprised 16 items, which formed 4 scales reflecting 4 aspects of attitude: cognitive (6 items), emotional-positive (3 items), emotional-negative (3 items), and motivational (4 items). We developed the structure of this scale based on exploratory and confirmatory factor analyses (unpublished data, 2017). Each item started with the statement "Usually, the elderly patients (65+) after a visit at my office…" followed by statements indicating older patients' ATH, such as "understand the nature and causes of their ailments," feel calmer," "have fears about their symptoms," "are going to participate in the treatment actively." GPs responded on a 7-point Likert scale (1=definitely no, 7=definitely yes). Each scale score was calculated as a mean value (the sum of item scores divided by the number of items in a scale). The possible scale score was between 1 and 7. Higher scores suggest a more active attitude in all aspects, except for the negative emotions. Reliability coefficients of PRACTA-ATH-D scales were good or very good and ranged from .88 to .93 before the intervention and from .88 to .94 after the intervention.

PRACTA-SE-D is a unidimensional scale consisting of 3 items that measure GPs' perception of older patients' self-efficacy related to health behavior changes: the fifth aspect of ATH included in the study. The statement "Usually, the elderly patients (65+) after a visit at my office…" was followed by items indicating their patients' sense of self-efficacy, such as "think they can influence how they'll feel in the future." Possible responses were from 1=definitely no to 7=definitely yes. The score was calculated as a mean value. Higher scores indicated greater self-efficacy. The reliability coefficients of the PRACTA-SE-D scale at time 1 and time 2 were .90.

Statistical Analysis

To compare the study groups with respect to descriptive statistics, we used a chi-square test (for frequencies) and analysis of variance (for interval and ratio scales; in cases where the assumption of variance homogeneity for analysis of variance was not met, we applied the Brown-Forsythe correction) [40].

In the final study group, there were no missing values at time 1, and at time 2 they were <0.5%. The Little missing completely at random test pointed to random missingness (P=.15).

To take into account possible changes occurring in the measured variables between time 1 and time 2, we calculated the indexes of change for each variable (as the time 2–time 1 difference). The index of change above zero indicated an increase in GPs' ratings (ie, GPs at time 2 rated older patients' expectations as more important, the patients' attitude as more active, and their own communication behavior as more frequent than those at time 1) and the index of change below zero demonstrated a decrease in these ratings between time 1 and time 2. Subsequently, we applied the indexes in the tested models as outcome variables.

In most cases, the Kolmogorov-Smirnov test showed that the outcome variables were not normally distributed (P<.05). Thus, to analyze the significance of differences between indexes of change in groups and to test interaction effects, we used the generalized linear model [41]. In the tables below, we present the results of the generalized linear model performed without an intercept (unstandardized B parameters equal to mean values of variables). We used pairwise comparisons with Bonferroni correction for multiple comparisons. We used the same type of analysis to compare the study groups in variables measured at time 1.

We used the following procedure of power calculation. Because our pilot study didn't have an interventional design, it was impossible to measure the standard deviations of indexes of change of the study variables. We performed a posteriori power calculation for the indexes of change of ATH scales (as key indicators of activation) and used it to verify adequacy of the sample size. We assumed an alpha of .05 and level of power of 80%. The required population size to detect effect sizes defined as a pairwise difference at the level of 0.50 and 0.25 (absolute value) were estimated. The analyses indicated that the size of the groups needed to detect an effect size estimated at 0.50 ranged from 30 to 40 per group depending on the ATH scale, and to detect an effect size estimated at 0.25, the required group size ranged from 118 to 158 per group (the only exception was the negative emotions scale, which needed 107 and 425 participants, respectively).

Results

Participants' Characteristics

Table 1 presents descriptive statistics for the e-learning, pdf article, and control groups. The analyzed groups differed with respect to some features of the facilities. In the e-learning group, there were more doctors working in privately owned facilities and in facilities where times of visits were scheduled individually for every patient. In the e-learning group, there were also fewer doctors working in facilities located in bigger towns (>100,000 residents) and in facilities where the average single visit lasted longer than 15 minutes. Doctors in the e-learning group worked in facilities where the average number of patients assigned to a single doctor was significantly lower than in control group facilities. As for factors concerning the doctors, the total number of working hours per week was significantly lower in the e-learning group than in the pdf article group.

Table 1. Descriptive statistics of the study groups.

Characteristics	Group			Test of differences	*P* value
	E-learning (n=42)	Pdf article (n=89)	Control (n=94)		
Factor describing primary care facilities					
Location (no. of inhabitants), n (%)					
<100,000[a]	13 (35)	27 (33)	26 (39)	$\chi^2_4=27.4$	<.001
>100,000	3 (8)	27 (33)	42 (45)		
Capital city (Warsaw)	21 (57)	27 (33)	14 (15)		
Organizational structure of facility, n (%)					
State owned	14 (37)	53 (60)	65 (69)	$\chi^2_2=11.7$	.003
Privately owned	24 (63)	36 (40)	29 (31)		
Visits system[b], n (%)					
Numbers	2 (53)	20 (23)	22 (24)	$\chi^2_4=11.7$	.02
Scheduled time	32 (84)	49 (55)	56 (62)		
Order of arrival	4 (11)	20 (23)	13 (14)		
Average time of visit, n (%)					
<15 minutes	17 (49)	27 (31)	22 (26)	$\chi^2_2=6.1$	.048
>15 minutes	18 (51.4)	59 (68.6)	64 (74.4)		
Average no. of patients per doctor in facility, mean (SD)	1444 (425)	1681 (672)	1754 (791)	$\text{B-F}_{2,215}=3.33$[c]	.04
No. of doctors working in facility, mean (SD)	5.45 (3.12)	4.79 (2.93)	5.41 (3.38)	$F_{2,218}=1.08$	.34
Factors describing doctors					
Age in years, mean (SD)	49.56 (11.56)	49.44 (11.35)	50.39 (13.16)	$F_{2,218}=0.15$	.86
Sex, n (%)					
Female	36 (86)	62 (70)	62 (66)	$\chi^2_2=5.7$	.06
Male	6 (14)	27 (30)	32 (34)		
Marital status, n (%)					
Single	4 (10)	12 (14)	8 (9)	$\chi^2_6=3.2$	.79
Married	33 (79)	65 (73)	77 (82)		
Divorced/widowed	5 (12)	12 (14)	9 (10)		
Seniority, mean (SD)	23.90 (12.13)	23.57 (11.99)	23.87 (13.15)	$F_{2,220}=0.02$	.98
Hours weekly in facility, mean (SD)	33.89 (9.48)	32.72 (10.82)	31.34 (9.89)	$F_{2,219}=0.98$	.38
Hours weekly overall, mean (SD)	39.53 (11.01)	45.36 (15.23)	41.54 (13.21)	$F_{2,219}=3.07$	.048
Training in geriatrics[d], n (%)					
None	28 (67)	49 (55)	49 (52)	$\chi^2_4=6.6$	.16
Single	12 (29)	30 (34)	27 (29)		
Multiple	2 (5)	10 (11)	18 (13)		
Percentage of older patients[e], n (%)					
<25%	3 (7)	14 (16)	10 (11)	$\chi^2_6=3.7$	.72
25%-50%	19 (45)	32 (36)	33 (35)		

Characteristics	Group			Test of differences	*P* value
	E-learning (n=42)	Pdf article (n=89)	Control (n=94)		
51%-75%	16 (38)	36 (40)	40 (43)		
≥75%	4 (10)	7 (8)	11 (12)		
Specialization, n (%)					
Internal medicine	15 (37)	28 (34)	46 (55)	χ^2_6=11.1	.09
Family medicine	15 (37)	32 (39)	24 (29)		
2 specializations[f]	9 (22)	13 (16)	9 (11)		
Others	2 (5)	10 (12)	5 (6)		

[a] This category includes both rural areas and small towns.

[b] Appointment systems were (1) numbers, which informed patients about their place in a queue to see a doctor but not about the time of their appointment (usually it forced patients to come in advance and wait for a long time not to miss their appointment), (2) scheduled time (the patient was informed about the exact time of their appointment), (3) order of arrival (patients were free to choose the time of their appointment but there was no control over patient flow).

[c] B-F: Brown-Forsythe test.

[d] Training in geriatrics encompassed any form of a postgraduate course.

[e] Doctors' ratings of percentage of older patients (age ≥65 years) among their patients last year.

[f] Two specializations when at least one was internal medicine or family medicine.

We examined differences between the group taking part in the baseline questionnaire and the final study group in terms of all controlled variables concerning facilities and doctors. We found that doctors in the final study group, in comparison with the dropout group, worked in facilities in which there were significantly fewer doctors ($F_{1,467}$=14.18, *P*<.001; mean 5.15, SD 3.16 and mean 6.42, SD 3.96, respectively) and doctors working in privately owned facilities dominated in this group (χ^2_1=11.7, *P*<.001). Moreover, doctors in the final study group reported a much higher number of working hours in the given facility (Brown-Forsythe $\text{test}_{1,482}$=8.68, *P*=.003; mean 32.35, SD 10.20 and mean 29.39, SD 11.84, respectively) and a much lower number of doctors from this group declared multiple training in geriatrics (χ^2_2=6.5, *P*=.04).

Level of Outcome Variables at Time 1

We compared the study groups with respect to initial levels of all variables used to calculate the indexes of change. Significant differences were found in GPs' perception of older patients' expectations for treatment explanation (Wald χ^2=10.6, *P*=.005) and health advice (Wald χ^2=6.2, *P*=.046). In both cases, the pdf article group scored lower than the control group (*P*=.001 and *P*=.01, respectively). Further differences were found in 3 aspects of ATH: positive emotions (Wald χ^2=8.1, *P*=.017), negative emotions (Wald χ^2=9.6, *P*=.008), and motivation (Wald χ^2=6.2, *P*=.045). The e-learning group scored lower than the control group in perception of older patients' positive emotions (*P*=.006) and motivation (*P*=.007), and higher in negative emotions (*P*=.004). On the motivational scale, the pdf article group scored lower than the control group as well (*P*=.046). Multimedia Appendix 5 shows trellis plots presenting time 1 and time 2 section and scale scores of each tool for each study group.

Effects of the PRACTA Intervention on GPs' Perception of Older Patients' Expectations

Table 2 presents means, standard errors, and comparative results of indexes of change in the GPs' perception of older patients' expectations in each study group, in the form of significance of Wald chi-square tests and pairwise comparisons.

Table 2. Indexes of change in general practitioners' perception of older patients' expectations by study group.

Variable	Group[a]			Wald χ^2[b] (*P* value)	Pairwise comparisons (*P* value)[c]		
	EL (n=42) mean (SE)	A (n=89) mean (SE)	C n=94) mean (SE)		EL–C	A–C	EL–A
1. Disease explanation							
	.71 (.165)	.15 (.131)	.01 (.108)	19.7 (<.001)	0.69 (.001)	0.14 (>.99)	0.55 (.03)
2. Treatment explanation							
	.33 (.172)	.11 (.121)	–.06 (.092)	4.9 (.18)	0.39 (.14)	0.17 (.83)	0.22 (.86)
3. Health advice							
	.23 (.221)	.10 (.136)	–.27 (.128)	6.0 (.11)	0.50 (.15)	0.37 (.15)	0.13 (>.99)
4. Quality of life							
	.17 (.323)	.07 (.179)	–.28 (.194)	2.5 (.45)	0.45 (.67)	0.35 (.55)	0.10 (>.99)
5. Rapport							
	–.01 (.099)	–.27 (.088)	–.15 (.070)	14.1 (.003)	0.13 (.77)	–0.13 (.76)	0.27[d] (.04)
6. Emotional support							
	–.29 (.140)	–.30 (.121)	.12 (.147)	11.2 (.01)	–0.41 (.13)	–0.42 (.08)	0.01 (>.99)

[a] Study groups were e-learning (EL), pdf article (A), and control (C).

[b] Wald chi-square test of the overall model.

[c] Pairwise comparisons with Bonferroni correction.

[d] Pairwise comparison significant only after use of least squares difference test.

The results presented in Table 2 demonstrate that the greatest differences between the groups regarded the index of change in GPs' perception of older patients' expectations for disease explanation. In the e-learning group the importance of this expectation increased more than in the control and pdf article groups. In relation to the GPs' perception of older patients' expectations for emotional support and rapport, the overall models also indicated significant main effects of the intervention but with no significant pairwise comparisons. In the case of expectation for emotional support, the pairwise differences only approached significance, indicating that in the pdf article group, the perception of the importance of older patients' expectations for emotional support decreased in comparison with the control group. In the case of expectations for rapport, applying the least squares difference test (a method less restrictive than the Bonferroni correction) revealed that the index of change in the e-learning group was significantly higher than in the pdf article group (*P*=.04). What was not significantly affected by the intervention was GPs' perception of older patients' expectations for treatment explanation, health advice, and quality-of-life improvement.

Effects of the PRACTA Intervention on GPs' Self-Assessed Communication With Older Patients

Table 3 presents means, standard errors, and the results of the comparison of indexes of change in communication behavior (global score and specific items) among GPs in each study group, in the form of significance of Wald chi-square tests and pairwise comparisons.

Table 3. Indexes of change in general practitioners' self-assessed communication[a] by study group.

Scale item	Group[b]			Wald χ^{2c} (*P* value)	Pairwise comparisons (*P* value)[d]		
	EL (n=42) mean (SE)	A (n=89) mean (SE)	C (n=94) mean (SE)		EL–C	A–C	EL–A
Global communication score							
	.19 (.126)	–.44 (.079)	–.09 (.074)	34.5 (<.001)	0.28 (.18)	–0.35 (.003)	0.63 (<.001)
1. I greet them in a kind manner.							
	.05 (.193)	–.21 (.101)	.06 (.110)	4.9 (.18)	–0.01 (>.99)	–0.28 (.19)	0.26 (.69)
2. I discuss with them the reason for a visit.							
	.02 (.180)	–.23 (.117)	–.31 (.121)	10.2 (.02)	0.33 (.38)	0.08 (>.99)	0.25 (.74)
3. I listen to them carefully.							
	.17 (.190)	–.34 (.094)	–.12 (.096)	15.1 (.002)	0.28 (.55)	–0.22 (.31)	0.50 (.053)
4. I show understanding of their problems.							
	.02 (.163)	–.42 (.109)	–.15 (.097)	16.9 (.001)	0.17 (>.99)	–0.27 (.20)	0.44 (.08)
5. I make sure I understood them correctly.							
	.14 (.211)	–.35 (.127)	.07 (.134)	8.3 (.04)	0.07 (>.99)	–0.42 (.07)	0.49 (.14)
6. I encourage them to ask questions.							
	.62 (.286)	–.18 (.165)	.04 (.168)	5.9 (.11)	0.58 (.25)	–0.22 (>.99)	0.80 (.046)
7. I answer all their questions.							
	–.07 (.163)	–.47 (.114)	–.42 (.120)	29.5 (<.001)	0.34 (.27)	–0.06 (>.99)	0.40 (.13)
8. I make sure they understood me correctly.							
	.07 (.166)	–.63 (.126)	–.13 (.125)	26.4 (<.001)	0.20 (>.99)	–0.50 (.01)	0.70 (.002)
9. I use language they can understand.							
	.00 (.135)	–.56 (.100)	–.11 (.108)	32.3 (<.001)	0.10 (>.99)	–0.48 (.006)	0.56 (.002)
10. I summarize topics we discussed.							
	.21 (.204)	–.47 (.133)	–.15 (.131)	15.0 (.002)	0.36 (.40)	–0.32 (.25)	0.67 (.01)
11. I inform them about the examination.							
	.07 (.229)	–.27 (.135)	.02 (.165)	4.1 (.25)	0.05 (>.99)	–0.29 (.52)	0.34 (.60)
12. I care about their comfort during the examination.							
	.21 (.181)	–.52 (.101)	–.28 (.125)	32.0 (<.001)	0.49 (.08)	–0.24 (.41)	0.73 (.001)
13. I provide as much time as they need for each part of the visit.							

Scale item	Group[b]			Wald χ^2[c] (*P* value)	Pairwise comparisons (*P* value)[d]		
	EL (n=42) mean (SE)	A (n=89) mean (SE)	C (n=94) mean (SE)		EL–C	A–C	EL–A
	.21 (.281)	–.51 (.183)	–.15 (.165)	9.0 (.03)	0.36 (.80)	–0.35 (.45)	0.72 (.10)
14. I explain treatment options available in their situation.							
	.36 (.220)	–.38 (.126)	–.13 (.109)	13.3 (.004)	0.48 (.15)	–0.25 (.38)	0.74 (.01)
15. I explain why they should comply with the recommendations.							
	.14 (.174)	–.40 (.109)	–.14 (.107)	16.2 (.001)	0.28 (.50)	–0.27 (.24)	0.55 (.02)
16. I make sure they’ll be able to comply with the recommendations.							
	.02 (.167)	–.40 (.134)	.11 (.130)	9.7 (.02)	–0.08 (>.99)	–0.51 (.02)	0.43 (.14)
17. I write down the main recommendations for them.							
	–.02 (.145)	–.43 (.103)	–.13 (.088)	19.4 (<.001)	0.10 (>.99)	–0.30 (.08)	0.40 (.07)
18. I discuss the plan of further treatment.							
	.29 (.182)	–.40 (.109)	–.06 (.114)	16.6 (.002)	0.35 (.31)	–0.34 (.09)	0.69 (.003)
19. I briefly summarize the entire visit.							
	.36 (.222)	–.46 (.140)	.05 (.148)	13.5 (.004)	0.30 (.77)	–0.51 (.04)	0.82 (.006)
20. I encourage them to participate in making decisions.							
	.64 (.164)	–.43 (.148)	.05 (.138)	23.8 (<.001)	0.59 (.02)	–0.48 (.053)	1.07 (<.001)
21. I give the opportunity to express their opinion.							
	.62 (.168)	–.48 (.145)	–.10 (.132)	25.2 (<.001)	0.71 (.002)	–0.38 (.14)	1.10 (<.001)
22. I take their opinion into account in making decisions.							
	.38 (.178)	–.66 (.146)	.11 (.145)	25.9 (<.001)	0.25 (.69)	–0.77 (.001)	1.04 (<.001)
23. I create an atmosphere that allows them to discuss intimate issues freely.							
	.21 (.163)	–.48 (.131)	.02 (.127)	16.6 (.001)	0.19 (>.99)	–0.50 (.01)	0.68 (.002)
24. I notice their feelings and accept them.							
	.07 (.170)	–.51 (.106)	.03 (.132)	23.2 (<.001)	0.04 (>.99)	–0.54 (.005)	0.58 (.01)
25. I ensure a good atmosphere during the entire visit.							
	.10 (.183)	–.66 (.121)	–.23 (.100)	35.8 (<.001)	0.33 (.35)	–0.43 (.02)	0.76 (.002)
26. I win their trust.							

Scale item	Group[b] EL (n=42) mean (SE)	A (n=89) mean (SE)	C (n=94) mean (SE)	Wald χ^2[c] (P value)	Pairwise comparisons (P value)[d] EL–C	A–C	EL–A
	.07 (.156)	–.53 (.119)	–.21 (.116)	23.2 (<.001)	0.28 (.43)	–0.31 (.18)	0.60 (.007)

[a] Assessed by the Promoting Active Aging Communication Scale-Doctors (PRACTA-CS-D) in response to questions about general practitioners' usual behavior with their older patients (≥65 years).

[b] Study groups were e-learning (EL), pdf article (A), and control (C).

[c] Wald chi-square test of the overall model.

[d] Pairwise comparisons with Bonferroni correction.

As Table 3 shows, the intervention had a significant impact on the global communication score. Importantly, the changes observed in the pdf article group and the e-learning group were in opposite directions, with a decrease in the pdf article group and an increase in the e-learning group. The index of change in the pdf article group was much lower than in the control and e-learning groups.

At the level of specific GPs communication behavior, between-group differences were significant in 20 of 26 analyzed cases (significant in the overall model and in the difference between at least 2 groups). The mean values of the indexes of change in the pdf article group and e-learning group demonstrated that changes were in opposite directions. In 19 cases, in the pdf article group indexes of change in communication with older patients were significantly different from those in the e-learning group, and in most cases (n=12) also in the control group. The most distinct differences between the e-learning and the pdf article groups (with the difference exceeding the level of 1) were in the following behaviors: encouraging older patients to participate in making decisions, giving them the opportunity to express their opinions, and taking their opinions into account in making decisions. It is worth noting that such differences exceeded the level of 0.5 in a further 13 items.

There were no between-group differences in only 5 cases: greeting older patients in a kind manner, discussing the reasons for their visit, answering all their questions, informing them about the examination, and providing as much time as they needed. In case of 1 item ("I encourage them to ask questions") the overall model was not significant, but a very large positive change in the e-learning group resulted in e-learning having a significant effect compared with the pdf article.

Effects of the PRACTA Intervention on GPs' Perception of Older Patients' ATH at the End of the Visit

Table 4 presents means, standard errors, and comparative results of indexes of change concerning GPs' perception of older patients' ATH at the end of the visit, in the form of significance of Wald chi-square tests and pairwise comparisons.

Table 4. Indexes of change in general practitioners' perception of older patients' attitude toward treatment and health by study group.

Variable	Group[a]			Wald χ^{2}[b] (*P* value)	Pairwise comparisons (*P* value)[c]		
	EL (n=42) mean (SE)	A (n=89) mean (SE)	C (n=94) mean (SE)		EL–C	A–C	EL–A
1. Cognitive aspect							
	.26 (.163)	–.13 (.120)	–.11 (.097)	5.02 (.17)	0.37 (.15)	–0.03 (>.99)	0.40 (.15)
2. Positive emotions							
	.19 (.145)	–.20 (.102)	–.17 (.098)	8.7 (.03)	0.36 (.11)	–0.02 (>.99)	0.39 (.08)
3. Negative emotions							
	–.54 (.267)	–.04 (.196)	–.23 (.190)	5.5 (.14)	–0.31 (>.99)	0.18 (>.99)	–0.50 (.41)
4. Motivational aspect							
	.35 (.175)	–.17 (.114)	–.18 (.107)	8.9 (.03)	0.53 (.03)	0.01 (>.99)	0.52 (.04)
5. Self-efficacy							
	.35 (.165)	.01 (.129)	–.05 (.104)	4.9 (.17)	0.41 (.11)	0.06 (>.99)	0.35 (.29)

[a] Study groups were e-learning (EL), pdf article (A), and control (C).

[b] Wald chi-square test of model effect.

[c] Pairwise comparison with Bonferroni correction.

The strongest between-group effects occurred in relation to changes in GPs' perception of older patients' motivation for active participation. There were significant differences between the e-learning group and the control and pdf article groups. Similar but weaker effect occurred in relation to GPs' perception of positive emotions demonstrated by older patients at the end of the visit. The difference between the e-learning and pdf article groups only approached significance. The intervention did not significantly affect GPs' perception of the following aspects of ATH: cognitive, negative emotions, and self-efficacy.

Moderators of Intervention Effects

We found effects of moderation in 2 organizational factors: the number of patients assigned to a single GP and facility's organizational structure.

We found group × number of patients per GP interactions in the following types of expectations: disease explanation (Wald χ^2=15.7, *P*<.001), treatment explanation (Wald χ^2=12.9, *P*=.002), quality-of-life improvement (Wald χ^2=13.1, *P*=.001), and health advice (Wald χ^2=5.7, *P*=.06). In the e-learning group, in contrast to the other groups, there was a positive relationship between the number of patients per GP and the GPs' perception of older patients' expectations for disease explanation. An increase in the number of patients per GP by 1000 increased the importance of older patients' expectations for disease explanation by an average of 0.95 (B=0.95, *P*=.04) and increased treatment explanation expectation by 1.05 (B=1.05, *P*=.001). At the same time, in the case of quality-of-life improvement, an increase in the number of patients per GP by 1000 resulted in a decrease in the outcome by 2.01 (B= –2.01, *P*=.01) in the e-learning group and by 1.38 (B= –1.38, *P*=.001) in the pdf article group. In the pdf article group, along with the increasing number of patients per GP by 1000, the perception of the importance of older patients' expectation for health advice decreased by 0.86 (B= –0.86, *P*=.02). We also found group × number of patients per GP interactions in relation to GPs' perception of 2 aspects of older patients' ATH: negative emotions (Wald χ^2=27.5, *P*<.001) and self-efficacy (Wald χ^2=9.0, *P*=.01). Contrary to the other groups, in the e-learning group, together with the increase by 1000 of the number of patients per GP, negative emotions declined by 3.2 (B= –3.2, *P*<.001) and perception of older patients' self-efficacy increased by 1.33 (B=1.33, *P*=.004).

Facility organizational structure differentiated the effects of the intervention in terms of GPs' perception of older patients' motivation for active participation (Wald χ^2=10.8, *P*=.004) and their self-efficacy (at the level of statistical trend: Wald χ^2=5.4, *P*=.07). Doctors in the e-learning group working in privately owned facilities reported an increase in older patients' motivation by 1.13 (B=1.13, *P*=.005) and in their self-efficacy by 0.85 (B=0.85, *P*=.02), whereas the changes in other intervention groups and type of facility were not significant. In the case of positive emotions, the effect of interaction was not statistically significant (Wald χ^2=4.6, *P*=.10), but in the e-learning group, doctors working in privately owned facilities scored higher by 0.77 (B=0.77, *P*=.04) than those in other the groups. Moreover, the effect of the statistical trend occurred in relation to GP's perception of the importance of older patients' expectations for disease explanation (Wald χ^2=5.8, *P*=.05). Doctors in the e-learning group working in privately owned facilities reported an increase in perception of such older patient needs by 0.83 (B=0.83, *P*=.02) compared with the other groups.

Discussion

Principal Results and Comparison With Prior Work

GPs' Perception of Older Patients' Expectations

The strongest effect in changes in GPs' perception of older patients' expectations for the medical encounter occurred in the e-learning group, in which there was a larger increase in their perception of the importance of explaining the patient's disease than in the control and pdf article groups. There were similar, though not statistically significant, increases in expectations for health advice and treatment explanation. Doctors in the e-learning group were more likely after than before the intervention to state that such needs were important to older patients. Such changes may help to create an image of older patients free from negative stereotypes [30]. For instance, the stereotype that older age means reduced cognitive abilities, including understanding and appropriate use of information about the disease, can lead to infantilization of the elderly, often manifested in the form of elderspeak [42] and moralizing attitudes about older patients [43]. The intervention had no effect on expectations for quality-of-life improvement. The question arises as to how strongly doctors believe that dealing with this aspect of older patients' lives belongs to the competence of a doctor [44].

Although the main effects of the intervention regarding expectations for rapport and emotional support were significant, differences between groups only approached significance. In the pdf article group, the doctors after completing the intervention stated that expectations for emotional support and rapport were not as important for older patients as they had thought before the intervention. These changes could have been due to the doctors' increased awareness of older patients' needs and, in consequence, to relativization of the meaning of the expectation for rapport and emotional support. There is a fairly common belief among physicians that older patients visit their GP solely for social reasons [31]. As a result, doctors may overestimate the importance of older patients' emotional and relational needs and focus less on the need for information and medical care [24].

Generally, the main impact of e-learning was the growing importance of older patients' cognitive expectations, especially their expectations for disease explanation, in comparison with the pdf article group. This would allow for only partial confirmation of hypothesis 1a. We formulated no directional hypotheses regarding expectations for rapport and social support, but the decrease we observed in the pdf article group appears to be slightly larger than that in the e-learning group.

GPs' Self-Assessed Communication With Older Patients

In terms of self-assessed communication, the 2 forms of intervention yielded different effects, with a substantial decrease in the pdf article group and a moderate increase in the e-learning group. Doctors in the pdf article group less frequently behaved in a desired manner after the intervention, both at the level of the global score and at the level of 20 of 26 specific behaviors. Desired GP behavior seems unlikely to have decreased as a result of the intervention. What might be likely, on the other hand, is that their assessment of their own behavior changed, as compared with the behavior described in the pdf article; doctors might have realized how their behavior differed from what was the most desirable. This would demonstrate doctors' growing self-criticism on reading the pdf article. It is worth noting that only 5 of the 26 described behaviors were not affected by the pdf article at all. Paradoxically, although the direction of change was negative, the intervention achieved the desired effect. In the e-learning group, 2 types of behavior increased in frequency. They both concerned actions aimed at activating older patients. This means that the pdf article contributed to GPs' critical self-assessment of their own communication behavior, but that e-learning would have a greater potential to enhance skills or GPs' subjective sense of their improved skills [34]. Although in the majority of cases the increase in self-assessed communication in the e-learning group was not significantly larger than in the control group, it was larger than in the pdf article group. We did not hypothesize that we would see such a decrease in self-assessed communication in the pdf article group, but hypothesis 1b may be partially confirmed.

GPs' Perception of Older Patients' ATH

The strongest effect on ATH was in the e-learning group, in GPs' perception of older patients' motivation to actively participate was significantly greater than in the control and pdf article groups. The second important result was the rise, compared with the pdf article group, in the level of positive emotions observed in older patients. Both aspects are the essence of successful aging [36]. It is worth noting that, contrary to the pdf article and control groups, in the e-learning group changes were positive in all aspects of ATH (apart from negative emotions). The results regarding changes in doctors' perception of older patients' attitudes as a result of participating in e-learning are consistent with hypothesis 1c, but the pdf article did not have any significant effect in this area.

Moderators of Intervention Effects

Among numerous factors considered as potential moderators of the intervention, only 2 organizational factors proved to be significant. In the e-learning group, having more patients assigned to the doctor increased the importance of expectations for disease and treatment explanation, increased the level of patient self-efficacy, and lowered the intensity of their negative emotions as perceived by GPs. This might mean that doctors with a greater number of patients achieved more of the desired effects of e-learning.

However, in both intervention groups, having more patients reduced the importance of expectations for quality-of-life improvement. Numerous factors are associated with the number of patients assigned to a doctor, and further research is required to determine which of them contributed most to the results. One can only speculate that, for example, higher remuneration and job satisfaction of doctors with a greater number of patients would foster greater involvement and motivation to expand their competences; on the other hand, more duties would result in cutting down the appointment time, reducing the focus on the patient, and disregarding their needs for improving quality of life.

Another factor that moderated the effects of the intervention was the facility organizational structure. Among doctors in the e-learning group working in privately owned facilities, the effect concerning perception of older patients as more motivated to actively participate was significantly higher than in the other groups. Furthermore, as shown by other effects, although weaker, it was this group who benefited from this type of intervention more than the others. This suggests that privately owned facilities create better conditions for e-learning to be effective, as they offer better organizational conditions, or GPs working in or running private medical services are more likely to have psychosocial dispositions such as openness to new experience or flexibility. Generally, the results confirm the validity of studies looking for potential moderators of intervention effects [33].

Limitations

This study had selection bias [32]. Although recruitment of facilities for the study was random, only one-fifth of managers approved and then only half of the invited doctors took part in the assessment at time 1. One can speculate that such approval was given only by managers who were not afraid of any evaluation and who welcomed research and innovations. Despite random allocation of facilities to the intervention, only some doctors participated in the interventions and the assessment at time 2 (attrition bias). More GPs in the final group than in the dropout group worked in privately owned facilities and in facilities employing a lower number of doctors. Moreover, doctors from these facilities declared more working hours in a given facility, but less training in geriatrics. This may indicate that GPs who had agreed to participate in the program enjoyed a better organizational setting and might be more dedicated to their facilities. They also seemed to be good prospects for training in older patient activation.

As a result of attrition, the study groups sizes were not fully satisfactory in relation to the results of power sample calculations (as they were insufficient to detect smaller differences). Moreover, the study groups were not balanced at entry, in terms of both some descriptive statistics and selected study variables. In the e-learning group more GPs worked in privately owned facilities located in the capital or small towns or villages, with more patient-customized appointment systems (scheduled individually) and fewer patients assigned to a single GP. These features seemed to be conducive to participation in the intervention, but in this study they did not turn out to be sufficient. As for outcome variables at time 1, the e-learning group scored lower than the control group in perception of older patients' positive emotions and motivation, and higher in perception of negative emotions. These differences might partially explain why we observed significant effects of e-learning in some of these aspects of ATH. On the other hand, the pdf article group scored lower than the control group in perception of older patients' expectations for treatment explanation and health advice, but with no impact on the effects of the pdf article in such areas.

However, the described recruitment process reflects the actual situation, and confirms extreme difficulties in conducting such a study under natural conditions and in implementing the intervention; these limitations were also reported in other studies [32-34]. This means that the results of this study should be interpreted with caution and can't be generalized to the entire population of GPs. Apart from the above-described limitations, there were also specific inclusion criteria, such as having a contract signed with the National Health Fund and the location in central Poland. The first criterion suggests that the results apply only to patients who did not cover the cost of medical service themselves, whereas the location criterion suggests that any conclusions should be restricted to GPs and older patients living in wealthier areas of the country.

Generally, the study group seemed to demonstrate specific features that might have been responsible for relatively small effect sizes of the intervention, as there were also effects of statistical trends in addition to statistically significant effects. As mentioned above, the sample sizes of the study groups were insufficient to detect smaller differences, and only those above the level of 0.5 were marked as significant.

It should be noted that all outcome variables were subjective. They encompassed GPs' perception of older patients and GPs' perception of their own performance; no performance, however, was objectively measured. Thus, the analysis of older patients' attitudes in the context of doctors' attitude change would be a valuable complement to the results.

A separate analysis is needed to discover the reasons for such a low proportion of doctors in e-learning. The reasons can be divided into 2 groups depending on the time of making the decision: before or after entering the intervention. As for reasons, which might matter before the start, the e-learning requirement to use more advanced technology seems to be a disincentive. There was no age difference between the study groups. The participants' average age was about 50 years. Thus, low computer literacy and a fear of advanced technology can't be ruled out as deterrents to starting the intervention. Participation in the pdf article intervention, where the use of a computer was only required to read the article, was higher. Of 208 doctors in the pdf article and control groups who received the opportunity to participate in PRACTA e-learning at the end of the time 2 assessment, 117 declared a desire to participate but ultimately only 7 doctors logged on to the platform. Perhaps GPs would like to participate in such training but reasons such as lack of time, fatigue, and the need to focus on urgent tasks prevent them from participating.

On the other hand, 11 (21%) of 53 GPs who only logged on and a further 10 (19%) who completed no more than 3 modules decided to drop out of the study after starting the e-learning intervention. The question arises as to whether the course in the form of a multimedia game demanding active participation from the learner was adequate for such a group of GPs. Additionally, each module took about an hour to complete, and the range of subjects and cognitive load could have been excessive for some

GPs. At the same time, the majority of GPs who had started the e-learning (71%) completed all 5 modules.

Another important issue was the extent of correspondence between the e-learning and pdf article interventions. We assumed that both forms of the intervention were identical in terms of messages directed to learners but were distinct from each other in their delivery methods. The main objective of the pdf article was to describe active aging and GPs' ways of enhancing active aging, whereas the objective of the e-learning was not only to describe it but also to demonstrate (to model) how to do it. Based on theories of social learning, knowledge is important in changing attitudes and behavior, but is insufficient on its own. Only combined with training and creating new habits does knowledge increase the amplitude of such change [38]. Therefore, after having read the pdf article, GPs were expected to know what to do (what is recommended), but they did not receive any training in how to do it (as a consequence, the change in outcomes was smaller), whereas after e-learning they were expected to be equipped not only with knowledge but also with skills (as a consequence, the change in outcomes was larger). However, it is very likely that the methods used in the e-learning (eg, case study, video) might at least slightly have changed the message contained in the text material (number of details, associations, practical implications). Therefore, it seems of great importance to precede implementation of the intervention with a pilot study to verify content consistency of both forms and to correct any discrepancies. Doing so would give deeper insights into factors responsible for the achieved results.

The above limitations clearly demonstrate the need to replicate the study in larger groups. Further research is required to verify content consistency of both forms of the intervention, but also the relationships between the studied variables and their impact on the achieved effects of the intervention. A valuable complement to the results would be to analyze the patients' attitudes in the context of doctors' behavioral change. This will be the next step in our ongoing project.

Conclusions

As far as expectations are concerned, the main impact of e-learning was the growing importance of older patients' cognitive expectations, especially the expectation for disease explanation. In terms of self-assessed communication, the 2 forms of intervention yielded different effects, with a substantial decrease in the pdf article group and a moderate increase in the e-learning group. Despite the negative direction of changes in the pdf article group, such a result may be perceived as a desired effect of the intervention, as it reflects a critical self-assessment by doctors of their own communication behavior. In terms of the ATH, the strongest effect concerned the e-learning group, in which there was a significant increase in GPs' perception of older patients' motivation for active participation and positive emotions compared with the control and pdf article groups. Among many factors moderating the effects of intervention, 2 factors of organizational character proved to be significant: the number of patients assigned to a single GP and the facility organizational structure. In the e-learning group, having more patients assigned to the doctor increased the importance of expectations for disease and treatment explanation, and of perceiving higher levels of self-efficacy and a lower intensity of negative emotions among older patients. In both intervention groups, having more patients reduced the importance of the expectation for quality-of-life improvement. In the e-learning group of doctors working in privately owned facilities, the effect of perceiving older patients as being more motivated to actively participate was significantly higher than in the other groups.

The results demonstrate the suitability of the 2 methods, but in other areas and under different conditions. The key benefit of the pdf article intervention was doctors' growing reflection on their limitations in terms of communication skills, whereas e-learning was more effective in changing the perception of older patients' proactive attitude, especially among GPs working in privately owned facilities and having a higher number of assigned patients. Although we did not achieve all the expected effects of the PRACTA intervention, both its forms seem promising in terms of growing competencies of GPs in communication with and activation older patients.

Acknowledgments

Funding for the study was provided by Norway Grants in the Polish-Norwegian Research Program (Pol-Nor/200856/34/2013).

Authors' Contributions

DW was project leader. All authors contributed to the conception and design of the project. DW conceptualized research problems and the manuscript structure, carried out the statistical analyses, and wrote the manuscript. JC participated in researching the literature and writing the section pertaining to methods. ML coordinated the preparation of the educational intervention. MA and MJ coordinated data collection. MR, ML, GH, and GAE contributed to the conception of this study and reviewed manuscript drafts.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Contents of the Promoting Active Aging (PRACTA) intervention.

[PDF File (Adobe PDF File), 118KB - jmir_v19i2e45_app1.pdf]

Multimedia Appendix 2

Promoting Active Aging (PRACTA) Patients Expectations Scale - Doctors (PRACTA-PES-D).

[PDF File (Adobe PDF File), 297KB - jmir_v19i2e45_app2.pdf]

Multimedia Appendix 3

Promoting Active Aging (PRACTA) Communication Scale - Doctors (PRACTA-CS-D).

[PDF File (Adobe PDF File), 309KB - jmir_v19i2e45_app3.pdf]

Multimedia Appendix 4

Promoting Active Aging (PRACTA) Attitude Toward Treatment and Health Scale - Doctors (PRACTA-ATH-D) and PRACTA Self-Efficacy Scale - Doctors (PRACTA-SE-D).

[PDF File (Adobe PDF File), 296KB - jmir_v19i2e45_app4.pdf]

Multimedia Appendix 5

Trellis plots of time 1 and time 2 section and scale scores for e-learning, pdf article, and control groups.

[PDF File (Adobe PDF File), 149KB - jmir_v19i2e45_app5.pdf]

References

1. Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: the challenges ahead. Lancet 2009 Oct 3;374(9696):1196-1208 [FREE Full text] [doi: 10.1016/S0140-6736(09)61460-4] [Medline: 19801098]
2. Mallers MH, Claver M, Lares LA. Perceived control in the lives of older adults: the influence of Langer and Rodin's work on gerontological theory, policy, and practice. Gerontologist 2014 Feb;54(1):67-74. [doi: 10.1093/geront/gnt051] [Medline: 23723436]
3. World Health Organization. Good health adds life to years: global brief for World Health Day 2012. Geneva, Switzerland: WHO; 2012. URL: http://apps.who.int/iris/bitstream/10665/70853/1/WHO_DCO_WHD_2012.2_eng.pdf [accessed 2017-02-03] [WebCite Cache ID 6o0S14WLA]
4. Davis MM, Bond LA, Howard A, Sarkisian CA. Primary care clinician expectations regarding aging. Gerontologist 2011 Dec;51(6):856-866 [FREE Full text] [doi: 10.1093/geront/gnr017] [Medline: 21430129]
5. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. J Gen Intern Med 2012 May;27(5):520-526 [FREE Full text] [doi: 10.1007/s11606-011-1931-2] [Medline: 22127797]
6. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013 Feb;32(2):207-214. [doi: 10.1377/hlthaff.2012.1061] [Medline: 23381511]
7. Ory MG, Peck BM, Browning C, Forjuoh SN. Lifestyle discussions during doctor-older patient interactions: the role of time in the medical encounter. MedGenMed 2007 Dec 04;9(4):48. [Medline: 18311398]
8. Singer S, Shortell SM. Implementing accountable care organizations: ten potential mistakes and how to learn from them. JAMA 2011 Aug 17;306(7):758-759. [doi: 10.1001/jama.2011.1180] [Medline: 21828308]
9. Frosch DL, Rincon D, Ochoa S, Mangione CM. Activating seniors to improve chronic disease care: results from a pilot intervention study. J Am Geriatr Soc 2010 Aug;58(8):1496-1503 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.02980.x] [Medline: 20662953]
10. Gerber LM, Barrón Y, Mongoven J, McDonald M, Henriquez E, Andreopoulos E, et al. Activation among chronically ill older adults with complex medical needs: challenges to supporting effective self-management. J Ambul Care Manage 2011;34(3):292-303. [doi: 10.1097/JAC.0b013e31821c63b1] [Medline: 21673530]
11. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in the Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). BMC Public Health 2012;12:577 [FREE Full text] [doi: 10.1186/1471-2458-12-577] [Medline: 22849664]
12. Maio G, Haddock G. The Psychology of Attitudes and Attitude Change. 2nd edition. London, UK: SAGE; 2010.
13. Wrześniewski K. The development of a scale for assessing attitudes toward illness in patients experiencing a myocardial infarction. Soc Sci Med Med Psychol Med Sociol 1980 Mar;14A(2):127-132. [Medline: 7209611]
14. Dattilo J, Martire L, Gottschall J, Weybright E. A pilot study of an intervention designed to promote walking, balance, and self-efficacy in older adults with fear of falling. Educ Gerontol 2013 Sep 06;40(1):26-39. [doi: 10.1080/03601277.2013.768067]

15. Bowling A, Rowe G, Lambert N, Waddington M, Mahtani KR, Kenten C, et al. The measurement of patients' expectations for health care: a review and psychometric testing of a measure of patients' expectations. Health Technol Assess 2012 Jul;16(30):i-xii, 1 [FREE Full text] [doi: 10.3310/hta16300] [Medline: 22747798]
16. Zolnierek KB, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. Med Care 2009 Aug;47(8):826-834 [FREE Full text] [doi: 10.1097/MLR.0b013e31819a5acc] [Medline: 19584762]
17. Williams SL, Haskard KB, DiMatteo MR. The therapeutic effects of the physician-older patient relationship: effective communication with vulnerable older patients. Clin Interv Aging 2007;2(3):453-467 [FREE Full text] [Medline: 18044195]
18. Valori R, Woloshynowych M, Bellenger N, Aluvihare V, Salmon P. The Patient Requests Form: a way of measuring what patients want from their general practitioner. J Psychosom Res 1996 Jan;40(1):87-94. [Medline: 8730648]
19. Aldwin CM, Spiro A, Park CL. Health, behavior, and optimal aging: a life span developmental perspective. In: Birren JE, Schaie KW, editors. Handbook of the Psychology of Aging. San Diego, CA: Elsevier Academic Press; 2006:85-104.
20. Tarn DM, Young HN, Craig BM. Development of the patient approach and views toward healthcare communication (PAV-COM) measure among older adults. BMC Health Serv Res 2012 Aug 30;12:289 [FREE Full text] [doi: 10.1186/1472-6963-12-289] [Medline: 22931537]
21. Parchman ML, Zeber JE, Palmer RF. Participatory decision making, patient activation, medication adherence, and intermediate clinical outcomes in type 2 diabetes: a STARNet study. Ann Fam Med 2010;8(5):410-417 [FREE Full text] [doi: 10.1370/afm.1161] [Medline: 20843882]
22. Peck BM. Age-related differences in doctor-patient interaction and patient satisfaction. Curr Gerontol Geriatr Res 2011;2011:137492 [FREE Full text] [doi: 10.1155/2011/137492] [Medline: 22007206]
23. Stewart M, Meredith L, Brown JB, Galajda J. The influence of older patient-physician communication on health and health-related outcomes. Clin Geriatr Med 2000 Feb;16(1):25-36, vii. [Medline: 10723615]
24. Delgado A, Andrés LL, de Dios Luna J, Gil N, Jiménez M, Puga A. Patient expectations are not always the same. J Epidemiol Community Health 2008 May;62(5):427-434. [doi: 10.1136/jech.2007.060095] [Medline: 18413456]
25. Palmore E. Ageism comes of age. J Gerontol B Psychol Sci Soc Sci 2015 Nov;70(6):873-875. [doi: 10.1093/geronb/gbv079] [Medline: 26362603]
26. Quadagno J. Aging and the Life Course: An Introduction to Social Gerontology. 4th edition. New York, NY: McGraw-Hill; 2008.
27. Scheidt RJ. The defense of my aging self: a report from the field. Gerontologist 2017 Feb;57(1):110-115. [doi: 10.1093/geront/gnw082] [Medline: 27106826]
28. Samra R, Griffiths A, Cox T, Conroy S, Gordon A, Gladman JRF. Medical students' and doctors' attitudes towards older patients and their care in hospital settings: a conceptualisation. Age Ageing 2015 Sep;44(5):776-783 [FREE Full text] [doi: 10.1093/ageing/afv082] [Medline: 26185282]
29. Hess T. Attitudes toward aging and their effects on behavior. In: Birren JE, Schaie KW, editors. Handbook of the Psychology of Aging. San Diego, CA: Elsevier Academic Press; 2006:379-405.
30. Meisner BA. Physicians’ attitudes toward aging, the aged, and the provision of geriatric care: a systematic narrative review. Crit Public Health 2012 Mar;22(1):61-72. [doi: 10.1080/09581596.2010.539592]
31. Ouchida KM, Lachs MS. Not for doctors only: ageism in healthcare. Generations 2015;39(3):46-57.
32. Sinclair PM, Kable A, Levett-Jones T, Booth D. The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: a systematic review. Int J Nurs Stud 2016 May;57:70-81. [doi: 10.1016/j.ijnurstu.2016.01.011] [Medline: 27045566]
33. Taveira-Gomes T, Ferreira P, Taveira-Gomes I, Severo M, Ferreira MA. What are we looking for in computer-based learning interventions in medical education? A systematic review. J Med Internet Res 2016 Aug 01;18(8):e204 [FREE Full text] [doi: 10.2196/jmir.5461] [Medline: 27480053]
34. Feng JY, Chang YT, Chang HY, Erdley WS, Lin CH, Chang YJ. Systematic review of effectiveness of situated e-learning on medical and nursing education. Worldviews Evid Based Nurs 2013 Aug;10(3):174-183. [doi: 10.1111/wvn.12005] [Medline: 23510119]
35. Riediger M, Li SC, Lindenberger U. Selection, optimization, and compensation as developmental mechanisms of adaptive resource allocation: review and preview. In: Birren JE, Schaie KW, editors. Handbook of the Psychology of Aging. San Diego, CA: Elsevier Academic Press; 2006:289-313.
36. Katz S, Calasanti T. Critical perspectives on successful aging: does it “appeal more than it illuminates”? Gerontologist 2015 Feb;55(1):26-33 [FREE Full text] [doi: 10.1093/geront/gnu027] [Medline: 24747713]
37. Rollnick S, Miller W, Butler CC. Motivational Interviewing in Health Care: Helping Patients Change Behavior. New York, NY: The Guilford Press; 2008.
38. Schwarzer R. Health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Appl Psychol 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
39. Underwood PW. Social support: the promise and the reality. In: Rice VH, editor. Handbook of Stress, Coping and Health: Implications for Nursing, Research, Theory and Practice. Thousand Oaks, CA: SAGE Publications; 2000:367-392.
40. Field A. Discovering Statistics Using SPSS. 3rd edition. London, UK: SAGE Publications; 2009:247-294.
41. Dobson AJ. An Introduction to Generalized Linear Models. 2nd edition. Boca Raton, FL: Chapman and Hall/CRC; 2002.

42. Williams K, Kemper S, Hummert ML. Enhancing communication with older adults: overcoming elderspeak. J Psychosoc Nurs Ment Health Serv 2005 May;43(5):12-16. [Medline: 15960030]
43. Higashi RT, Tillack AA, Steinman M, Harper M, Johnston CB. Elder care as “frustrating” and “boring”: understanding the persistence of negative attitudes toward older patients among physicians-in-training. J Aging Stud 2012 Dec;26(4):476-483. [doi: 10.1016/j.jaging.2012.06.007] [Medline: 22939544]
44. Shultz CG, Jimbo M. Decision aid use in primary care: an overview and theory-based framework. Fam Med 2015 Oct;47(9):679-692 [FREE Full text] [Medline: 26473560]

Abbreviations

ATH: attitude toward treatment and health
GP: general practitioner
PRACTA: Promoting Active Aging
PRACTA-ATH-D: Promoting Active Aging Attitude Toward Treatment and Health Scale-Doctors
PRACTA-CS-D: Promoting Active Aging Communication Scale-Doctors
PRACTA-PES-D: Promoting Active Aging Patients Expectations Scale-Doctors
PRACTA-SE-D: Promoting Active Aging Self-Efficacy Scale-Doctors
RQ: research question

Edited by G Eysenbach; submitted 05.11.16; peer-reviewed by T Taveira-Gomes, M Jimbo; comments to author 28.11.16; revised version received 07.01.17; accepted 29.01.17; published 22.02.17.

Please cite as:
Wlodarczyk D, Chylińska J, Lazarewicz M, Rzadkiewicz M, Jaworski M, Adamus M, Haugan G, Lillefjell M, Espnes GA
Enhancing Doctors’ Competencies in Communication With and Activation of Older Patients: The Promoting Active Aging (PRACTA) Computer-Based Intervention Study
J Med Internet Res 2017;19(2):e45
URL: http://www.jmir.org/2017/2/e45/
doi:10.2196/jmir.6948
PMID:28228370

Original Paper

Internet as a Source of Long-Term and Real-Time Professional, Psychological, and Nutritional Treatment: A Qualitative Case Study Among Former Israeli Soviet Union Immigrants

Anat Gesser-Edelsburg[1*], PhD; Svetlana Shalayeva[2*], BSc, MAN, RD

[1]University of Haifa Health and Risk Communication Research Center, School of Public Health, University of Haifa, Haifa, Israel
[2]School of Public Health, University of Haifa, Haifa, Israel
[*]all authors contributed equally

Corresponding Author:
Anat Gesser-Edelsburg, PhD
University of Haifa Health and Risk Communication Research Center
School of Public Health
University of Haifa
199 Aba Khoushy Ave. Mount Carmel
Haifa, 3498838
Israel
Phone: 972 544243530
Fax: 972 36322143
Email: ageser@univ.haifa.ac.il

Abstract

Background: The Internet is considered to be an effective source of health information and consultation for immigrants. Nutritional interventions for immigrants have become increasingly common over the past few decades. However, each population of immigrants has specific needs. Understanding the factors influencing the success of nutrition programs among immigrants requires an examination of their attitudes and perceptions, as well as their cultural values.

Objective: The purpose of this study was to examine perceptions of the Internet as a tool for long-term and “real-time” professional, psychological, and nutritional treatment for immigrants from the former Soviet Union who immigrated to Israel (IIFSU) from 1990 to 2012.

Methods: A sample of nutrition forum users (n=18) was interviewed and comments of 80 users were analyzed qualitatively in accordance with the grounded theory principles.

Results: The results show that IIFSU perceive the Internet as a platform for long-term and “real-time” dietary treatment and not just as an informative tool. IIFSU report benefits of online psychological support with professional dietary treatment. They attribute importance to cultural customization, which helps reduce barriers to intervention.

Conclusions: In light of the results, when formulating nutritional programs, it is essential to have a specific understanding of immigrants’ cultural characteristics and their patterns of Internet use concerning dietary care.

(J Med Internet Res 2017;19(2):e33) doi:10.2196/jmir.7130

KEYWORDS

long-term care; real-time systems; online systems, health psychology; nutrition therapy; qualitative research; former Soviet Union immigrants

Introduction

Internet as a Source of Consultation and Support for Immigrants

Personalized treatment with psychological support in nutritional intervention over the Internet increases the effectiveness of interventions [1,2]. Social media offer a platform for long-term and real-time support with professional consultation, rather than frontal meetings [3-5].

The Internet is considered one of the most effective tools for reducing medical gaps between immigrants and locals [6,7]. It is an important tool for overcoming immigrants’ feelings of estrangement and language barriers [8]. Research in different countries has shown that, compared with the majority population, immigrants generally have less access to health services and health information, mainly due to language and cultural barriers [9-11]. Subsequently, online health services offered by multicultural professionals serve as a tool for reducing communication barriers, thereby improving the quality and delivery of health services [12].

The use of Internet among immigrants is correlated with online health literacy. However, there are populations of immigrants, like Russian immigrants, who are well educated, but whose health literacy is low, and consequently, their knowledge of health and nutrition is poor [13,14]. This may present a challenge in a new country with a different culture and unfamiliar eating habits. Therefore, it is essential to mentor immigrants while considering their cultural needs.

Internet Intervention Programs and Real-Time Support

Continued use of nutrition interventions by health professionals is associated with successful results among patients [2,15]. Also, real-time support is essential to the success of the intervention [16].

Despite the benefits of the Internet as a platform for dietary treatment including professional and psychological support, few studies have focused on long-term and real-time online nutritional intervention. An interdisciplinary review [2] has shown that most of the studies in the field are randomized controlled trials with a maximum period of 13 months per intervention. Qualitative studies that examined attitudes of participants in nutritional interventions often included semistructured and indepth interviews [17,18] and focus groups [18].

Studies that were conducted among immigrants involved school children, adolescents, and young adults and their families [19]; overweight and obese and sick adults [20]; older adults [17]; pregnant women [21]; or males and females [22]. Some of the studies were conducted on bilingual health assistants [21,22]. Nearly all the studies considered first-generation immigrants and were small scale.

These interventions usually included a predefined plan prepared by the authors [23-25]. The participants did not always have the opportunity to ask questions or to obtain information about personal issues in real time. Long-term online nutritional treatment for Russian immigrant populations has not been studied. This study examines the use of Internet as a tool for long-term dietary counseling and psychological support for immigrants from the former Soviet Union in accordance with their cultural needs.

Characteristics of Russian Immigrants in Israel

IIFSU tend to have advanced degrees [26], but this is not always reflected in health literacy. Despite their high level of education, their nutritional knowledge and dietary intake are poor.

After immigration, IIFSU often see themselves as belonging to both the Russian and Israeli cultures, meaning they have 2 cultural identities [17]. After arriving in Israel, nutrition becomes more significant to them [13,14]. In Israel, eating habits are different from that in Russia. According to Lissak [27] and Biederman [28], in contrast with Soviet immigrants who arrived in Israel in the 1970s and sought to assimilate into the collective culture, the IIFSU showed a tendency toward acculturation [29] (eg, by opening Russian supermarkets, theaters, and bookstores). When immigrants rely on their originary ethnic group that speaks their language and abide by its customs and norms, including eating habits, such behavior increases their sense of security and facilitates gradual entry into the broader society [30-33]. Ben-Sira [34] argues that the IIFSU tend to maintain a diet high in fat and cholesterol based on eating habits from their country of origin. Despite being aware that such a diet can lead to obesity and associated health risks, only 40% of the IIFSU tend to choose healthy food. These eating habits of diets based mostly on bread and processed meats, along with other unhealthy lifestyles such as heavy smoking, high alcohol intake, and infrequent doctor visits for preventative care are all part of the life history and health experiences prior to the dissolution of the Soviet Union [35-40]. The first Israeli National Health and Nutrition Survey [41] showed that obesity was more prevalent among Soviet immigrants who arrived in the 1990s than among other ethnic groups in Israel. This study examines attitudes and perceptions toward the Internet as a tool for long-term personalized nutritional treatment and support for IIFSU. In this study, the nutrition counseling was provided by professional certified nutritionists whose public record of their credentials was presented either on a website, on their forum, or on a business webpage.

Objectives

The objective of this study was to examine perceptions and attitudes about the use of Internet as a tool for dietary interventions among immigrants from the IIFSU. Specific objectives were (1) to examine the use of Internet as a source of long-term and real-time dietary treatment; (2) to examine benefits of online professional dietary treatment with psychological support among immigrants; and (3) to examine how the cultural customization helps reduce resistance and barriers to compliance with dietary treatment.

Methods

Study Design and Analysis

The study combined semistructured, indepth interviews, and comments made by Russian forum users on social networks (Facebook, odnoklassniki, vkontakte).

Data Collection

All interviewees provided information about background variables including age, education, marital status and employment, country of origin, and year of immigration. Questions pertaining to study objectives covered the following issues: Internet as a source of information for health issues, Internet use by IIFSU, centrality in their lives, and lack of adherence in general and regarding nutritional care. The protocol included questions on interviewees’ perceptions on nutritional issues, about the Internet as a source of knowledge, and their experiences of nutritional therapy interventions via the Internet;

their barriers and dilemmas concerning nutritional therapy, influenced by cultural characteristics, online health literacy, and acclimatization in the country. Participants were asked to compare face-to-face and Internet-based nutritional therapy.

Summary of the Recruitment Population Study

Different subtypes of the IIFSU Internet users were selected: diverse ages, gender, socioeconomic status, and number of years in Israel (Table 1).

Recruitment was carried out through an open forum for nutritional counseling on social networks. All users received nutritional consultation and long-term treatment. Some were followed up from 1 month to several years. Eighteen indepth face-to-face interviews (between 40 and 60 min) were conducted at a time and place convenient to interviewees. Interviews were recorded and transcribed. Eighty users' comments from forums were analyzed using content analysis.

IIFSU, men and women aged 24-58 years, who immigrated to Israel from 1990 to 2012, most (14 out of 18) were found to hold BA degrees or higher and their socioeconomic statuses vary.

Table 1. Demographic data.

Interviewee	Gender	Age in years	Employment	Year of immigration	Years living in Israel	Age of immigration	Education
1	M	33	Hi-tech marketing manager	1993	22	12	BA
2	M	29	Computer science student	1999	16	14	BA student
3	M	40	Hi-tech engineer	1994	21	20	BA, MA student
4	F	25	Bookkeeper	1998	17	9	Vocational training
5	F	24	Dental assistant	2000	15	9	Vocational training
6	F	30	Hi-tech programmer	1991	24	6	BA
7	F	27	Social worker	1992	23	5	BA
8	F	53	Musician	1997	18	35	BA
9	F	40	Photographer	1998	17	24	Vocational training
10	F	31	Housewife	2012	3	29	BA
11	F	47	Saleswoman	2002	13	37	BA
12	M	56	Sports store manager	1990	25	31	MA
13	M	26	Construction worker	2000	15	11	High school
14	F	25	Architecture student	2003	12	13	BA student
15	M	58	Unemployed	1990	25	33	BA
16	F	55	Nurse	1998	17	38	BA
17	F	52	Factory worker	1991	24	28	BA
18	M	38	Lawyer	1993	22	16	BA

Data Analysis

The audiotape was transcribed as soon as possible after each interview. Transcripts were then checked against the recordings to ensure accuracy. We conducted a content analysis of additional 80 users' comments on nutritional forums. The analysis was conducted throughout the data collection process and the focus was on issues related to the research questions [42] and on themes that arose in the personal interviews.

The themes of the generic subjects that appeared in the content analyses of the users' comments and the semistructured protocols were analyzed. Interviews and users' comments were analyzed individually to identify key themes and subthemes. Data aggregation indicated a saturation point because there were repetitions of themes in the interviews and there was no need for more interviews [43]. The findings presented in the study are an integration of the forum users' comments and the issues that emerged during the interviews.

Validity and Reliability

We used the triangulation method including diverse data sources (tools). This approach uses cross-referencing data and validation. Use of multiple sources allows improved understanding, control, validity, and reliability of the findings. The study included personal interviews (semistructured protocol) and analysis of users' comments from Internet forums. For validation, the results were compared with findings in the literature. The comparison showed similar conclusions. Consequently, the study received a basis for its validity [44]. In addition, recognition of the limitations helped improve the quality and validity of the study [45]. At each stage, the researcher compared and brought into line the participants' views and the construction of those views by the researchers. To reinforce study reliability and credibility, different subtypes IIFSU were selected.

Ethical Considerations

Application was submitted to the Faculty of Social Welfare and Health Sciences Ethics Committee for research with human subjects at Haifa University and full ethical approval (no. 106/14) was granted.

Results

Main Subthemes

Over the course of the research, common themes arose for different subpopulations of IIFSU, along with a distinct variance between those researched. We divided the results into 3 subthemes that are as follows: (1) receiving long-term treatment and real-time dietary consultation; (2) professional dietary treatment with psychological support; and (3) cultural tailoring to the needs of IIFSU.

Receiving Long-Term Treatment and Real-Time Dietary Consultation

Most interviewees and most users' comments indicate that the Internet offers an opportunity to receive long-lasting dietary treatment. Also, there is the option of real-time consultation from professionals. Nutrition forums serve as a therapeutic tool as discussed in the following.

Long-Term Therapist-Patient–Relationship Through Online Counseling

In Israel, professional certified independent nutritionists provide online nutrition counseling. They promote their online services mainly through social media, and provide them through online video chat, Facebook, or their forum. Online nutrition counseling is not regulated by law. Eleven of 18 interviewees (7 women, 4 men) sought long-term dietary counseling, an unlimited number of consultations like in other online programs or in the traditional clinics. Patients state their need for customized therapeutic framework for effective changing of behavioral habits. It is a long-term process and this is why they need long-term follow-up and encouragement.

> *With the professional counseling I receive on the Internet, I feel they can mentor me and help me to change my eating habits step by step. This way I achieve my goals. It helps me to learn about new culture and food. This difficult process of changing habits takes time…So it is important to me to have follow-up and discipline, since I progress better like this… It is easier for me to deal with the problems with a therapist than alone...* [M, 29]

Real-Time Connection

Receiving counseling when difficulties or questions arise, which are not predetermined enables the patient to follow through on the therapist's recommendations. Twelve of 18 interviewees (8 women and 4 men, 25-35 years, with no difference in education or duration of residence in Israel), positively noted the option of a real-time relationship with their consultant via the Internet:

> *I keep having questions all the time, which I want answers to, and through the Web I can ask those questions immediately in real time. I want to ask the dietitian's opinion in real time, go over my experiences together, exchange experiences, write down what I'm going through,...I think that doing this in a clinical setting is problematic, since there is an accessibility issue and you have to reschedule appointments. Online I can address professionals any time and talk with them, which is impossible to do at the clinic.* [M, 33]

Professional Dietary Treatment With Psychological Support

Internet as a Place for Consulting Professionals

Thirteen of 18 interviewees (men and women, 24-58 years) regard the Internet as a venue for receiving dietary treatment and consultation, instead of frontal consultations with dietitians. Moreover, the saturation of dietary information on the Web is overwhelming and requires guidance. Russian users reveal that they only trust professionals. Users' comments revealed that they sought an authoritative source to assist in decision making. The study participants reported that they were looking for guidance from a licensed professional. Before they began the counseling, they reported that they checked the counselor's details on Google or that they asked to see professional certification to validate whatever licensing and degrees they claimed to have. According to the participants, a good nutritionist was one with professional training.

> *I improve my diet based on the professional consultation I get in forums. There are too many information and blogs about nutrition, but not many good professionals. It's critical for me to get a proper consultation which will influence my food intake.* [M, 29]

Internet as a Source of Psychological Support During the Treatment

Apart from the professional consultation, patients express a need for encouragement and motivation. Also mentoring and coaching helps change habits. Eight of 11 women stated they had used the Internet in situations of uncertainty. Through the connection which enables sharing difficulties and getting psychological support, the experience of treatment becomes more effective. The patient senses the therapists' support, when he needs their guidance, as one of the interviewees notes this:

> *To me, changing eating habits is like changing culture. It's changing your roots. It seems almost impossible, because my eating culture is a part of me…but I do it because it is important for me to improve my health, and it's important for me physically. From a psychological standpoint, it is very difficult. It takes a lot of time and professional support who can encourage, explain, and give us the appropriate tools to deal with temptation and pressure.* [M, 29]

Cultural Tailoring to the Needs of IIFSU

Being immigrants influences the compliance with dietary treatment for several reasons:

Language

Eleven of 18 interviewees stated that speaking Russian is important in dietary treatment. Those who chose to receive counseling in Russian stressed that it was easier as they conduct their personal lives in Russian. They see themselves as part of a Russian community in Israel.

> *I prefer someone who speaks Russian, but who lives in Israel, someone like me… So that we will have a better understanding…Although I am fine with Hebrew, it is better speaking with someone who knows your native language to make matters clearer, it makes it easier for me.* [F, 27]

Segmentation of the findings regarding language choice for treatment was observed among users from different age groups. Both the younger (aged 20-40) and older (aged 40-56) users noted a preference for counseling in Russian. A difference of opinion was also observed among those who immigrated to Israel as children (ages 2-14), as well as among those who immigrated after age 15, and had been living in Israel between 1 and 10 years. They all shared a common characteristic: patients whose daily routines were in Russian preferred a Russian-speaking therapist. That means the phenomenon has less to do with the age and amount of time since immigrating, and more to do with the role of Russian in their daily lives.

Mentality and Culture During Treatment

Nine of 18 patients expressed the need for a therapist who was also a Russian immigrant. This fostered mutual understanding and a sense of identification. Furthermore, it helps cope with difficulties in a new culture.

> *I prefer that the therapist be Russian...Not only speaking the language, but also knowing what Russian culture is, what a Russian upbringing is...I think it's the mentality, I mean that they will be someone like me who lives in Israel and knows the Russian mentality. That is why I think that a Russian dietitian understands me better. When I say a Russian lifestyle with a Russian mentality – it is because I speak the Russian language, eat Russian foods and take part in Russian culture and understanding. It's hard to explain, but it's like dating an Israeli. It's fun and nice but we're not on the same wavelength, we don't fully understand each other.* [F, 30]

Discussion

Principal Findings

IIFSU look for professional consultation with psychological support that takes into account their cultural identity. The IIFSU interviewed tend to be more educated than the general population. This might explain the extensive use of the Internet as an informative and therapeutic instrument. However, these data contradict findings in the literature, in which immigrant populations of disadvantaged groups (elderly, poor, and chronically ill), use the Internet less, and do so less rationally [15,46,47]. All the users in this study stated that the Internet was effective, empowering, and helpful for dietary issues. Despite their high level of education, a gap in health literacy remains [48]. The population interviewed tends to have unhealthy eating habits, reflecting the eating habits in their native country and the complexities of immigration. The study participants reported on their eating habits to the interviewer. The reports expose eating habits that include a diet high in fat, sweets, and juices, with little consumption of vegetables and water. The Russian interviewees sought dietary support in terms of both health and culture. This is in keeping with the importance placed on “cultural sensitivity” in the literature. Cultural sensitivity entails awareness of cultural similarity and difference, which influence worldviews, values and beliefs, learning processes, and behavior [49]. Differences in culture must therefore be considered at every stage in dietary intervention. This may guarantee the efficacy of the treatment offered to Internet users [50].

According to the literature, publishing dietary and medical content in the immigrants' language and making it culturally appropriate helps carry messages of health [15,46]. Most users in this study claimed that speaking Russian was a crucial factor in creating trust. However, it appears that a preference for Russian as a language for dietary treatment depends on the parameters. This study is the first to assess the differences among the Russian immigrants who prefer Russian for their online treatment. The variance in the distribution of preferences did not depend on age group or number of years since immigrating; rather, those whose daily routines were in Russian preferred a Russian-speaking therapist.

The Russian interviewees stressed the importance of the therapist's mentality and knowledge of both Russian and Israeli cultures. They need to be counseled by a Russian immigrant to Israel. The professional literature supports these findings. One of the measures for diminishing health literacy gaps between the 2 population strata such as Russian immigrants and the permanent residents of Israel is the communication that stems from the patient's dominant culture [51]. Online health services offered by multicultural professionals are a valuable tool for reducing communication barriers [12] and can increase the compliance with dietary treatment. This may be why an online therapist can replace frontal treatment at a clinic.

In this study, participants stressed that they preferred treatment on the Internet because it enabled the creation of a long-term relationship between the therapist and the patient. Furthermore, long-term online dietary intervention allows for a real-time relationship and facilitates compliance. Thus, the patient could receive continual reinforcement and maintain motivation.

Most studies conducted on intervention programs were limited by time, where professional counseling was offered at predetermined points at the beginning, middle, and end [23,25,52]. Based on our findings, users valued counseling during crisis and periods of success, which are not predetermined.

Russian immigrants in the study turned to a dietary treatment which includes not only professional consultation, but also psychological support. Studies show that Internet users look for nutritional information from evidence-based and easily recognizable sources [53]. In addition, users need psychological support and interventional guidance [54,55]. Providing professional consultation with coaching methods and psychological support show high compliance and usability among the IIFSU in this study.

In this study, this finding was particularly notable among young users (aged 25-35 years), regardless of difference in education or time spent in Israel, with a majority of women over men. In the professional literature, several explanations emerged for our study's findings. A Russian study [56], which performed an extensive examination of the behavioral patterns of more than 16 million users of the popular Russian social network (My.Mail.Ru), discovered that women were more active on the Internet than men. In addition, women were willing to disclose more details about themselves and share personal information, such as their physical appearance, hobbies, occupation, family status, and more. Similar findings have been noted among social network users around the world [57,58].

Creating professional and emotional support during the intervention that considers the cultural needs of patients fosters higher health care quality in cross-cultural situations [59].

Limitations

The limitations of this study are that as it is a qualitative study, it does not include a representative sample of the study population. In addition, the next step should be further validation through randomized controlled trials and implementation.

In our study, we did not compare the efficacy of the nutritional counseling online versus frontal nutritional counseling, as our focus was on the perceptions of the participants regarding the Internet as a tool for long-term and "real-time" professional, psychological, and nutritional treatment. Over the course of the study, some of the participants reported weight loss and an improvement in clinical indicators. However, as noted, we did not gather quantitative clinical data that would enable a comparison with one-on-one nutritional counseling. Notwithstanding this, in the literature there are studies that indicate positive results when using online nutritional counseling, which have been found to lead to a significant improvement in nutrition and to a healthier lifestyle [60], weight loss, rise in consumption of fruits and vegetables, a decrease in consumption rate of fats and sugar, and a decrease in calorie consumption [54]. Other studies did not find a significant difference in BMI and the BMI z-score between the experimental group and the control group [61], nor did they find significant differences in serum, blood pressure, anthropometry, social support, and cholesterol [62].

Given that the field of online nutritional counseling is still a relatively new field, we think that more studies are needed in order to assess the effectiveness of online counseling as compared with on-one-one nutritional counseling. We can also assume that in the future, technological advances will improve the efficacy of online counseling, whether by incorporating virtual smart agents, and augmented and virtual reality, or by an increased presence of peers who will empower the patients, leading to more effective counseling.

Conclusions

This study is an examination of the perceptions of IIFSU regarding the Internet as a tool of nutritional treatment. It included long-term dietary treatment, real-time counseling, and nutritional information with psychological support. All the interventions were carried out with the immigrants' cultural needs in mind. Our study is the first to examine unique cultural perceptions and beliefs, which affect the rate of response to dietary treatment among a minority population.

Conflicts of Interest

None declared.

References

1. Ashman AM, Collins CE, Brown LJ, Rae KM, Rollo ME. A brief tool to assess image-based dietary records and guide nutrition counselling among pregnant women: an evaluation. JMIR Mhealth Uhealth 2016 Nov 04;4(4):e123 [FREE Full text] [doi: 10.2196/mhealth.6469] [Medline: 27815234]
2. Enwald HP, Huotari ML. Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review. J Med Internet Res 2010 Jun 28;12(2):e24 [FREE Full text] [doi: 10.2196/jmir.1409] [Medline: 20584698]
3. Cobb NK, Jacobs MA, Saul J, Wileyto EP, Graham AL. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomised controlled trial study protocol. BMJ Open 2014 Jan;4(1):e004089 [FREE Full text] [doi: 10.1136/bmjopen-2013-004089] [Medline: 24448847]
4. Kernot J, Olds T, Lewis LK, Maher C. Effectiveness of a facebook-delivered physical activity intervention for post-partum women: a randomized controlled trial protocol. BMC Public Health 2013 May 29;13:518 [FREE Full text] [doi: 10.1186/1471-2458-13-518] [Medline: 23714411]
5. Patrick K, Marshall SJ, Davila EP, Kolodziejczyk JK, Fowler JH, Calfas KJ, et al. Design and implementation of a randomized controlled social and mobile weight loss trial for young adults (project SMART). Contemp Clin Trials 2014 Jan;37(1):10-18 [FREE Full text] [doi: 10.1016/j.cct.2013.11.001] [Medline: 24215774]
6. Lee SK, Sulaiman-Hill CM, Thompson SC. Providing health information for culturally and linguistically diverse women: priorities and preferences of new migrants and refugees. Health Promot J Austr 2013 Aug;24(2):98-103. [doi: 10.1071/HE12919] [Medline: 24168735]

7. Tanaka M, Strong C, Lee S, Juon H. Influence of information sources on hepatitis B screening behavior and relevant psychosocial factors among Asian immigrants. J Immigr Minor Health 2013 Aug;15(4):779-787 [FREE Full text] [doi: 10.1007/s10903-012-9753-9] [Medline: 23238580]
8. Mesch G, Mano R, Tsamir J. Minority status and health information search: a test of the social diversification hypothesis. Soc Sci Med 2012 Sep;75(5):854-858. [doi: 10.1016/j.socscimed.2012.03.024] [Medline: 22633160]
9. Brzoska P, Razum O. [Prevention among migrants--problems in health care provision and suggested solutions illustrated for the field of medical rehabilitation]. Dtsch Med Wochenschr 2014 Sep;139(38):1895-1897. [doi: 10.1055/s-0034-1387238] [Medline: 25203550]
10. Khan NA, Saboor HT, Qayyum Z, Khan I, Habib Z, Waheed HT. Barriers to accessing the German health-care system for Pakistani immigrants in Berlin, Germany: a qualitative exploratory study. The Lancet 2013 Oct;382:18. [doi: 10.1016/S0140-6736(13)62179-0]
11. Spallek J, Zeeb H, Razum O. Prevention among immigrants: the example of Germany. BMC Public Health 2010 Feb 24;10:92 [FREE Full text] [doi: 10.1186/1471-2458-10-92] [Medline: 20181232]
12. Shpilko I. Russian-American health care: bridging the communication gap between physicians and patients. Patient Educ Couns 2006 Dec;64(1-3):331-341. [doi: 10.1016/j.pec.2006.03.014] [Medline: 16859861]
13. Imamura F, Micha R, Khatibzadeh S, Fahimi S, Shi P, Powles J, Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (NutriCoDE). Dietary quality among men and women in 187 countries in 1990 and 2010: a systematic assessment. Lancet Glob Health 2015 Mar;3(3):e132-e142 [FREE Full text] [doi: 10.1016/S2214-109X(14)70381-X] [Medline: 25701991]
14. Lissitsa LS, Peres Y. New immigrants and old timers: Identity and interrelations - research findings. In: Epstein AD, Fedorchenko AV, editors. Mass Migration and Its Impact on the Israeli Society. Moscow and Jerusalem: Institute of Israeli and Middle Eastern Studies - Hebrew University of Jerusalem - Open University of Israel; (in Russian); 2000:244-278.
15. Moore M, Bias RG, Prentice K, Fletcher R, Vaughn T. Web usability testing with a Hispanic medically underserved population. J Med Libr Assoc 2009 Apr;97(2):114-121 [FREE Full text] [doi: 10.3163/1536-5050.97.2.008] [Medline: 19404502]
16. Stroetmann K, Kubitschke L, Robinson S, Stroetmann V, Cullen K, McDaid D. Policy brief 13: How can telehealth help in the provision of integrated care? In: Health Systems and Policy Analysis. Copenhagen: World Health Organization; 2010.
17. Khvorostianov N, Elias N, Nimrod G. 'Without it I am nothing': The internet in the lives of older immigrants. New Media & Soc 2011 Nov 11;14(4):583-599. [doi: 10.1177/1461444811421599]
18. Yehle KS, Chen M, Plake KS, Yi JS, Mobley AR. A qualitative analysis of coronary heart disease patient views of dietary adherence and web-based and mobile-based nutrition tools. J Cardiopulm Rehabil Prev 2012;32(4):203-209 [FREE Full text] [doi: 10.1097/HCR.0b013e31825b4e6a] [Medline: 22760245]
19. Chen JL, Weiss S, Heyman MB, Cooper B, Lustig RH. The efficacy of the web-based childhood obesity prevention program in Chinese American adolescents (Web ABC study). J Adolesc Health 2011 Aug;49(2):148-154 [FREE Full text] [doi: 10.1016/j.jadohealth.2010.11.243] [Medline: 21783046]
20. Saha S, Leijon M, Gerdtham U, Sundquist K, Sundquist J, Arvidsson D, et al. A culturally adapted lifestyle intervention addressing a Middle Eastern immigrant population at risk of diabetes, the MEDIM (impact of Migration and Ethnicity on Diabetes In Malmö): study protocol for a randomized controlled trial. Trials 2013 Sep 03;14:279 [FREE Full text] [doi: 10.1186/1745-6215-14-279] [Medline: 24006857]
21. Samkange-Zeeb F, Ernst SA, Klein-Ellinghaus F, Brand T, Reeske-Behrens A, Plumbaum T, et al. Assessing the acceptability and usability of an Internet-based intelligent health assistant developed for use among Turkish migrants: results of a study conducted in Bremen, Germany. Int J Environ Res Public Health 2015 Dec 03;12(12):15339-15351 [FREE Full text] [doi: 10.3390/ijerph121214987] [Medline: 26633455]
22. Sanou D, O'Reilly E, Ngnie-Teta I, Batal M, Mondain N, Andrew C, et al. Acculturation and nutritional health of immigrants in Canada: a scoping review. J Immigr Minor Health 2014 Feb;16(1):24-34 [FREE Full text] [doi: 10.1007/s10903-013-9823-7] [Medline: 23595263]
23. Kattelmann KK, White AA, Greene GW, Byrd-Bredbenner C, Hoerr SL, Horacek TM, et al. Development of Young Adults Eating and Active for Health (YEAH) internet-based intervention via a community-based participatory research model. J Nutr Educ Behav 2014 Mar;46(2):S10-S25. [doi: 10.1016/j.jneb.2013.11.006] [Medline: 24456834]
24. McCambridge J, Bendtsen M, Karlsson N, White IR, Bendtsen P. Alcohol assessment & feedback by e-mail for university student hazardous and harmful drinkers: study protocol for the AMADEUS-2 randomised controlled trial. BMC Public Health 2013;13:949 [FREE Full text] [doi: 10.1186/1471-2458-13-949] [Medline: 24456668]
25. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013 Mar 28;28(3):CD008776. [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
26. Rozenbaum-Tamari Y, Damian N. First 5 absorption years of Soviet Union immigrants (1990-1995). In: Gindin R, Rozenbaum-Tamari Y, editors. Immigrant Absorption Israel - Research Abstracts. Jerusalem: Ministry of Immigrant Absorption, Research and Planing Department; 1996:51-56.

27. Lissak M. The immigrants from the Soviet Union between seclusion and integrating. In: Lissak M, editor. Israel Towards the Year 2000. Jerusalem: Magnes; 1996:1-24.
28. Biederman P. Social Distance of Young Newcomers from the Former USSR from Israeli Society. Haifa: School of Social Work, University of Haifa; 1997.
29. Aaronson N. Absorption of immigrant's students from Soviet Union in the mirror of the reference group model. Haifa: University of Haifa; 1994.
30. Kozulin A, Wanger A. Immigration without adaptation: The psychological world of immigrants. In: Prital D, editor. The Jews of the Soviet Union in Transition. Jerusalem: The Public Committee for the Soviet Union Jews; 1995:183-192.
31. Leshem E. The Israel public's attitudes toward the new immigrants of the 1990s. Immigration to Israel: Sociological Perspectives 1994;3:164-188.
32. Mirsky J, Prawer L. To immigrate as an adolescent. In: Horowitz T, editor. Children of Perestroika in Israel. Langham: University Press of America; 1999.
33. Rosenbaum Y, Farber H. Social absorption of immigrants. In: Loewenberg M, K ra us M, editors. Absorption and Welfare in Israel - Reader. Ramat Gan: Bar Ilan University; 1979:15-46.
34. Ben-Sira Z. Nutrition knowledge and habits among immigrants from Ethiopia and Soviet Union. Jerusalem: Adar; 1993.
35. Benisovich SV, King AC. Meaning and knowledge of health among older adult immigrants from Russia: a phenomenological study. Health Educ Res 2003 Apr;18(2):135-144 [FREE Full text] [doi: 10.1093/her/18.2.135] [Medline: 12729173]
36. Duncan L, Simmons M. Health practices among Russian and Ukrainian immigrants. J Community Health Nurs 1996;13(2):129-137. [doi: 10.1207/s15327655jchn1302_6] [Medline: 8764387]
37. Hosler AS, Melnik TA, Spence MM. Diabetes and its related risk factors among Russian-speaking immigrants in New York State. Ethn Dis 2004;14(3):372-377. [Medline: 15328938]
38. Ivanov LL, Buck K. Health care utilization patterns of Russian-speaking immigrant women across age groups. J Immigr Minor Health 2002;4(1):17-27. [doi: 10.1023/A:1013055110631] [Medline: 16228751]
39. Ivanov LL, Hu J, Pokhis K, Roth W. Preventive health care practices of former Soviet union immigrant women in Germany and the United States. Home Health Care Manag Pract 2010 Dec;22(7):485-491. [doi: 10.1177/1084822310370945]
40. Mehler PS, Scott JY, Pines I, Gifford N, Biggerstaff S, Hiatt W. Russian immigrant cardiovascular risk assessment. J Health Care Poor Underserved 2001 May;12(2):224-235. [doi: 10.1353/hpu.2010] [Medline: 11370189]
41. Keinan-Boker L, Noyman N, Chinich A, Green M, Nitzan-Kaluski D. Overweight and obesity prevalence in Israel: findings of the first national health and nutrition survey (MABAT). Isr Med Assoc J 2005 Apr;7(4):219-223 [FREE Full text] [Medline: 15847200]
42. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. 4th ed. Thousand Oaks, CA: Sage Publications; 2009.
43. Shkedi A. Words of Meaning: Qualitative Research - Theory and Practice. Tel-Aviv: Ramot - Tel-Aviv University; 2003.
44. Patton MQ. Qualitative research & evaluation method. 4th ed. Thousand Oaks, CA: Sage Publications; 2015.
45. Guillemin M, Gillam L. Ethics, reflexivity, and “ethically important moments” in research. Qual Inq 2004 Apr;10(2):261-280. [doi: 10.1177/1077800403262360]
46. Selsky C, Luta G, Noone A, Huerta EE, Mandelblatt JS. Internet access and online cancer information seeking among Latino immigrants from safety net clinics. J Health Commun 2013;18(1):58-70 [FREE Full text] [doi: 10.1080/10810730.2012.688248] [Medline: 23066874]
47. Wang MP, Wang X, Lam TH, Viswanath K, Chan SS. Health information seeking partially mediated the association between socioeconomic status and self-rated health among Hong Kong Chinese. PLoS One 2013;8(12):e82720 [FREE Full text] [doi: 10.1371/journal.pone.0082720] [Medline: 24349347]
48. Gesser-Edelsburg A, Endevelt R, Zemach M, Tirosh-Kamienchick Y. Food consumption and nutritional labeling among immigrants to Israel from the former Soviet Union. J Immigr Minor Health 2015 Apr;17(2):459-466. [doi: 10.1007/s10903-013-9885-6] [Medline: 23955168]
49. Lopez-De Fede A, Bowman RP, Ewing T, Hanna J, Stafford Robinson J. Building Culture Bridges. Bloomington, IN: National Educational Service; 1997.
50. Becker ER, Roblin DW. Translating primary care practice climate into patient activation: the role of patient trust in physician. Med Care 2008 Aug;46(8):795-805. [doi: 10.1097/MLR.0b013e31817919c0] [Medline: 18665059]
51. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health Literacy: a Prescription to End Confusion. Washington, DC: National Academies Press; 2004.
52. Harmsen IA, Doorman GG, Mollema L, Ruiter A, Kok G, de Melker HE. Parental information-seeking behaviour in childhood vaccinations. BMC Public Health 2013 Dec 21;13:1219 [FREE Full text] [doi: 10.1186/1471-2458-13-1219] [Medline: 24358990]
53. van Veen MR, Beijer S, Adriaans AM, Vogel-Boezeman J, Kampman E. Development of a website providing evidence-based information about nutrition and cancer: fighting fiction and supporting facts online. JMIR Res Protoc 2015 Sep 08;4(3):e110 [FREE Full text] [doi: 10.2196/resprot.4757] [Medline: 26350824]

54. Anderson-Bill ES, Winett RA, Wojcik JR, Winett SG. Web-based guide to health: relationship of theoretical variables to change in physical activity, nutrition and weight at 16-months. J Med Internet Res 2011 Mar 04;13(1):e27 [FREE Full text] [doi: 10.2196/jmir.1614] [Medline: 21447470]
55. Love B, Thompson CM, Crook B, Donovan-Kicken E. Work and “mass personal” communication as means of navigating nutrition and exercise concerns in an online cancer community. J Med Internet Res 2013 May 31;15(5):e102 [FREE Full text] [doi: 10.2196/jmir.2594] [Medline: 23728365]
56. Kisilevich S, Ang CS, Last M. Large-scale analysis of self-disclosure patterns among online social networks users: a Russian context. Knowl Inf Syst 2011 Sep 25;32(3):609-628. [doi: 10.1007/s10115-011-0443-z]
57. Christofides E, Muise A, Desmarais S. Information disclosure and control on Facebook: are they two sides of the same coin or two different processes? Cyberpsychol Behav 2009 Jun;12(3):341-345. [doi: 10.1089/cpb.2008.0226] [Medline: 19250020]
58. Henderson S, Gilding M. ‘I’ve Never Clicked this Much with Anyone in My Life’: trust and hyperpersonal communication in online friendships. New Media & Soc 2004 Aug;6(4):487-506. [doi: 10.1177/146144804044331]
59. Saha S, Beach MC, Cooper LA. Patient centeredness, cultural competence and healthcare quality. J Natl Med Assoc 2008 Nov;100(11):1275-1285. [doi: 10.1016/S0027-9684(15)31505-4] [Medline: 19024223]
60. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth platform for pregnancy on nutrition and lifestyle of the reproductive population: a survey. JMIR Mhealth Uhealth 2016 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
61. 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]
62. Verheijden M, Bakx JC, Akkermans R, van den Hoogen H, Godwin NM, Rosser W, et al. Web-based targeted nutrition counselling and social support for patients at increased cardiovascular risk in general practice: randomized controlled trial. J Med Internet Res 2004 Dec 16;6(4):e44 [FREE Full text] [doi: 10.2196/jmir.6.4.e44] [Medline: 15631968]

Edited by G Eysenbach; submitted 08.12.16; peer-reviewed by R Endevelt, K Kattelmann; comments to author 18.01.17; revised version received 22.01.17; accepted 22.01.17; published 03.02.17.

Please cite as:
Gesser-Edelsburg A, Shalayeva S
Internet as a Source of Long-Term and Real-Time Professional, Psychological, and Nutritional Treatment: A Qualitative Case Study Among Former Israeli Soviet Union Immigrants
J Med Internet Res 2017;19(2):e33
URL: http://www.jmir.org/2017/2/e33/
doi:10.2196/jmir.7130
PMID:28159729

Original Paper

Demographic and Indication-Specific Characteristics Have Limited Association With Social Network Engagement: Evidence From 24,954 Members of Four Health Care Support Groups

Trevor van Mierlo[1,2*], MScCH, MBA, GEMBA, MSc BMR; Xinlong Li[3*], MSc; Douglas Hyatt[3*], PhD; Andrew T Ching[3*], PhD

[1]Research Associate, Henley Business School, University of Reading, Henley-on-Thames, United Kingdom

[2]Evolution Health Systems Inc, Toronto, ON, Canada

[3]Rotman School of Managment, University of Toronto, Toronto, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Trevor van Mierlo, MScCH, MBA, GEMBA, MSc BMR
Research Associate
Henley Business School
University of Reading
Greenlands
Henley-on-Thames, RG93AU
United Kingdom
Phone: 44 1491 571454
Fax: 44 1491 571454
Email: tvanmierlo@evolutionhs.com

Abstract

Background: Digital health social networks (DHSNs) are widespread, and the consensus is that they contribute to wellness by offering social support and knowledge sharing. The success of a DHSN is based on the number of participants and their consistent creation of externalities through the generation of new content. To promote network growth, it would be helpful to identify characteristics of superusers or actors who create value by generating positive network externalities.

Objective: The aim of the study was to investigate the feasibility of developing predictive models that identify potential superusers in real time. This study examined associations between posting behavior, 4 demographic variables, and 20 indication-specific variables.

Methods: Data were extracted from the custom structured query language (SQL) databases of 4 digital health behavior change interventions with DHSNs. Of these, 2 were designed to assist in the treatment of addictions (problem drinking and smoking cessation), and 2 for mental health (depressive disorder, panic disorder). To analyze posting behavior, 10 models were developed, and negative binomial regressions were conducted to examine associations between number of posts, and demographic and indication-specific variables.

Results: The DHSNs varied in number of days active (3658-5210), number of registrants (5049-52,396), number of actors (1085-8452), and number of posts (16,231-521,997). In the sample, all 10 models had low R^2 values (.013-.086) with limited statistically significant demographic and indication-specific variables.

Conclusions: Very few variables were associated with social network engagement. Although some variables were statistically significant, they did not appear to be practically significant. Based on the large number of study participants, variation in DHSN theme, and extensive time-period, we did not find strong evidence that demographic characteristics or indication severity sufficiently explain the variability in number of posts per actor. Researchers should investigate alternative models that identify superusers or other individuals who create social network externalities.

(J Med Internet Res 2017;19(2):e40) doi:10.2196/jmir.6330

KEYWORDS

econometric models; social networking; social support; self-help groups; data mining; Internet; regression analysis; forecasting; superusers

Introduction

Background

Digital health social networks (DHSNs), otherwise known as discussion forums or peer-to-peer support groups, are in abundance [1-8]. Although the efficacy of these networks is still being evaluated, the consensus is that social support and knowledge sharing increase patient education, enhance self-management, and decrease burden on existing health services [9-16].

In an era of increasing health costs [17,18], an aging population [19-22], and an annual US $300 billion adherence problem [23-26], DHSNs are beginning to play an important role in improving the delivery of North American health services [27,28].

As we increasingly rely on technology to help us look after our health, management science is playing a greater role in using data to measure efficiencies [29-31]. In the case of DHSNs, analysis is now turning to mechanisms that drive growth, help attain sustainability, and generate positive network externalities.

Research on Social Network Structure, Growth, and Sustainability

As a discipline, social network theory (SNT) maps social capital and the strength of relationships in networks. Within a network, nodes are individual actors, and ties are the relationships between nodes. For decades, disciplines such as economics, political science, public health, marketing, and finance have analyzed real world relationships within networks of actors [32-37]. These studies typically leverage graph theory, sociograms, or stochastic models to examine relationships [38-40].

Recently, SNT has shifted toward the topology of scale-free networks. This stream of research investigates whether network growth is random, if networks evolve, follow encoded and organized principles [41-46], and if taxonomies of actors naturally exist [47-51].

Three Fundamentals of Digital Health Social Networks

In the context of this study, actors are DHSN registrants who have created, at minimum, 1 post. From this perspective, 3 fundamental principles guide network growth.

The first is the network's total number of posts. In most DHSNs, actor posts remain on the network, and each new post adds to the quantitative size and value of the community. Whether actors passively read, actively respond to, or agree or disagree with new content, the quantitative value of the network n increased with each new post by $n+1$. In management and economics literature this is referred to as positive network externalities [52].

Second is the number of actors in the network. If a network contains x actors, potential connections between actors is $x(x-1)$. The greater the number of actors, the greater the potential for network expansion and the generation of new externalities. This has been illustrated in the study of networks in demand-side economies, where the value of a product or service is directly related to the number of others who use it [53,54].

Third, the mathematical relationship between these 2 quantities (positive network externalities and number of actors) represents a power law [55-57]. Marketing experts have observed this phenomenon and have intuitively referred to it as the 1% rule or the 90-9-1 principle [58,59]. Both concepts are related to the Pareto principle [60], and applied empirically, they have shown to be intrinsic to social network structure [61-63].

Monitoring nodes and ties, and monitoring topologies are important considerations for those who manage social networks. However, these tasks are retrospective as they examine a network's past state. Methods to drive future growth and promote individual agency are required. As the creation of externalities governs the success of a network, it would be helpful to profile actors who create value by generating externalities [64].

The Interventions

The 4 interventions in this study [65-68] contained self-guided interactive behavior change treatment programs based on state-of-the-art best practice, and have been examined extensively in the literature [69-83]. A component of each of the interventions is a DHSN moderated by trained and paid employees. All posts are reviewed and approved by a moderator, and any post that does not address the indication is permanently removed. Moderators can also instantaneously communicate with all actors. Table 1 outlines each program's theoretical constructs and evidence base.

Table 1. Theoretical constructs and evidence-base.

Theoretical construct	Problem drinking	Depressive disorder	Panic disorder	Smoking cessation
Brief intervention [84]	X	X	X	X
Cognitive behavioral therapy [85]		X	X	
Gamification [86]	X	X	X	X
Harm reduction [87]	X			X
Health belief model [88]	X	X	X	X
Motivational interviewing [89]	X	X	X	X
Normative feedback [90]	X			X
Social cognitive theory [91]	X	X	X	X
Structured relapse prevention [92]	X			
Targeting and tailoring [93]		X	X	
Transtheoretical model [94]				X

Table 2 outlines intervention launch dates, data acquisition dates, number of registrants, number of actors, total posts, and number of subjects used in analysis from their intervention DHSN inception until December 31, 2015.

Table 2. Four social networks.

Social network	Social network launch date	Data acquisition date	Number of days active	Number of subjects registered in program	Number of actors, n (%)	Number of actor posts[a]	Number of subjects in analysis, n (%)[a]
Problem drinking	Dec 26, 2005	Dec 31, 2015	3658	5049	1085 (21.49)	16,231	4784 (94.75)
Depressive disorder	Feb 6, 2003	Dec 31, 2015	4712	11,675	2065 (17.69)	20,516	1958 (16.77)
Panic disorder	January 23, 2002	Dec 31, 2015	591	9783	3579 (36.58)	61,743	6151 (62.87)
Smoking cessation	Sep 26, 2001	Dec 31, 2015	5210	52,396	8452 (16.13)	521,997	12,061 (23.01)
Total	n/a[b]	n/a	18,671	78,903	15,181 (19.24)	620,487	25,178 (31.91)
Mean	n/a	n/a	4688	19,726	3795 (19.24)	155,122	6239 (31.63)

[a] Moderator posts removed.

[b] n/a: not applicable.

Data Collected at Registration

Demographic characteristics (age, gender, highest level of education obtained, current occupation), and indication-specific details (Table 3) were collected at registration. Program registration and participation were free; however, consenting to the use of personal data for research purposes was a requirement.

Table 3. Indication-specific data collected at registration.

Intervention	Indication-specific data	Measurement
Problem drinking	Average drinks per day	Drop-down menu 0-30+
	Program goal: cut down, stop, unsure	Likert scale
Depressive disorder	Depression rating over past 2 weeks	Likert scale 0-10
	Level of distress over past 2 weeks	Likert scale 0-10
	Level of interference over past 2 weeks	Likert scale 0-10
	Tried cognitive behavior therapy in the past	Yes or no
	Currently being treated	Yes or no
	Using program with health care professional	Yes or no
Panic disorder	Number of attacks over past 2 weeks	Drop-down menu 0-51+
	Average fear rating during attack	Likert scale 0-10
	Attack interference with average daily life	Drop-down menu 0-4
	Attack causing avoidance	Drop-down menu 0-4
	Tried cognitive behavior therapy in the past	Yes/No
	Use of program with health care professional	Yes/No
Smoking cessation	Smoking patterns: ≥ 1 cigarette per day, occasional smoker, recently quit	Drop-down menu
	Last cigarette: >24 hours, <24 hours	Radio button
	Cigarettes per day	Drop-down menu 0-100+
	Total years smoked	Drop-down menu 0-75+
	Minutes to first cigarette: >60, 31-60, 6-31, ≤5	Drop-down menu
	Past year quit attempts > 24 hours	Drop-down menu 0-10+
	Number of cohabitant smokers	Drop-down menu 0-10+
	Fagerstrom dependency score (very low, low, moderate, high, very high)	Internal calculation

Objective

As a first step in profiling actors based on characteristics, and to investigate the feasibility of developing predictive models that identify superusers in real time, the objective of this study was to examine the association between number of posts and actor demographic and indication-specific variables inputted at registration.

Methods

Sample

Data were extracted from the custom SQL DHSN databases of the 4 digital health interventions. As they contained full data sets, samples totaling 24,954 registrants and 3285 actors were used in the analysis (Table 4).

Table 4. Sample size.

Intervention	Sample size	Sample size actors	Sample size posts
Problem drinking	4484	884	12,914
Depressive disorder	1958	206	3190
Panic disorder	6151	585	18,921
Smoking cessation	12,061	1610	90,894
Total sample	24,954	3285	125,919

Regression Models

A total of 5 models were developed to explore whether posting behavior was associated with demographics characteristics and indication-specific severity amongst all registrants (Table 5).

Table 5. Regression models for all subjects.

Model	Equation
1	$\mathit{ProblemDrinkingPostsAllRegistrants} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{DrinksPerDay} + \beta_6\mathit{Goal} +$
2	$\mathit{DepressiveDisorderPostsAllRegistrants} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{Rating} + \beta_6\mathit{Distress} + \beta_7\mathit{Interference} + \beta_8\mathit{CBT} + \beta_9\mathit{Treated} + \beta_{10}\mathit{Professional} +$
3	$\mathit{PanicDisorderPostsAllRegistrants} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{Attacks} + \beta_6\mathit{Fear} + \beta_7\mathit{Interference} + \beta_8\mathit{Avoidance} + \beta_9\mathit{CBT} + \beta_{10}\mathit{Professional} +$
4	$\mathit{SmokingCessationPostsAllRegistrants} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{Patterns} + \beta_6\mathit{LastCigarette} + \beta_7\mathit{CigarettesPerDay} + \beta_8\mathit{YearsSmoked} + \beta_9\mathit{FirstCigarette} + \beta_{10}\mathit{PastQuits} + \beta_{11}\mathit{CohabitantSmokers} + \beta_{12}\mathit{FagerstromScore} +$
5	$\mathit{TotalPostsAllRegistrants} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} +$

Another 5 additional regression models were developed to explore whether posting behavior was associated with demographics characteristics and indication-severity amongst actors (Table 6).

Table 6. Regression models for actors.

Model	Equation
6	$\mathit{ProblemDrinkingPostsActors} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{DrinksperDay} + \beta_6\mathit{Goal} +$
7	$\mathit{DepressiveDisorderPostsActors} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{Rating} + \beta_6\mathit{Distress} + \beta_7\mathit{Interference} + \beta_8\mathit{CBT} + \beta_9\mathit{Treated} + \beta_{10}\mathit{Professional} +$
8	$\mathit{PanicDisorderPostsActors} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{Attacks} + \beta_6\mathit{Fear} + \beta_7\mathit{Interference} + \beta_8\mathit{Avoidance} + \beta_9\mathit{CBT} + \beta_{10}\mathit{Professional} +$
9	$\mathit{SmokingCessationPostsActors} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} + \beta_5\mathit{Patterns} + \beta_6\mathit{Last\ Cigarette} + \beta_7\mathit{CigarettesPerDay} + \beta_8\mathit{YearsSmoked} + \beta_9\mathit{FirstCigarette} + \beta_{10}\mathit{PastQuits} + \beta_{11}\mathit{CohabitantSmokers} + \beta_{12}\mathit{FagerstromScore} +$
10	$\mathit{TotalPostsActors} = \beta_0 + \beta_1\mathit{Age} + \beta_2\mathit{Gender} + \beta_3\mathit{Education} + \beta_4\mathit{Occupation} +$

Dummy variables were created for categorical data, with 1 dummy variable excluded during regressions. Analyses were performed with Stata version 13 (Stata Corp LLP, College Station, TX, USA).

As outlined in previous research conducted on the 4 DHSNs, the number of posts per actor is right skewed, indicating the presence of a power law [44]. Negative binomial regression was employed as the method of analysis for 3 reasons. First, the dependent variable in our model, number of observations, is counted as integers only. Second, negative binomial regression can capture the skewness of the data. Third, Poisson distribution requires the mean and the variance of the model to be identical and in each of the models, the hypothesis of equidispersion is rejected.

Ethics

All data collection policies and procedures adhered to international privacy guidelines [95-97] and were in accordance with the Helsinki Declaration of 1975, as revised in 2008 [98]. The study was consistent with the University Research Ethics Committee procedures at Henley Business School, University of Reading, and was exempt from full review.

Results

R-Squared Values

All 5 models had low R^2 values (see Table 7 and Multimedia Appendix 1).

Regression Analysis: Demographic Variables

A total of 4 independent demographic variables were included in each of the 10 models (Table 8).

In 9 of the models, age was positively and significantly associated with number of posts (beta range =.13-.4). This means that as age of registrants increased, number of posts increased marginally.

Education was positively and significantly associated to the number of posts in 6 models (beta range =.082-.315). This means that within these 6 models, number of posts increases by less than 1 with every unit increase in education category.

Gender was negatively and significantly associated number of posts in 4 models (beta range =−.766 to −.272). This means that within these 4 models, number of posts decreased by less than 1 with male registrants.

Registrants had the option of selecting from 1 of 12 occupations. Compared with registrants who indicated that they were full-time students, occupation was positively associated with number of posts in 14 cases (beta range =.377-5.301), and negatively associated with number of posts in 19 cases (beta range =-2.609 to -.587).

The variable *occupation not listed* was selected with the greatest frequency 60% (6/10), and was positively and significantly associated to the number of posts in 4 of these 6 models (beta range =.488-.703), but negatively and significantly associated to the number of posts in 2 of these 6 models (beta range =−1.314 to −.945).

Table 7. R^2 values for ten models.

Model	1	2	3	4	5	6	7	8	9	10
R^2	0.016	0.013	0.02	0.043	0.026	0.027	0.018	0.061	0.086	0.031

Table 8. Statistically significant demographic independent variables (all models).

Independent variable		Model 1 beta (*P* value)	Model 2 beta (*P* value)	Model 3 beta (*P* value)	Model 4 beta (*P* value)	Model 5 beta (*P* value)	Model 6 beta (*P* value)	Model 7 beta (*P* value)	Model 8 beta (*P* value)	Model 9 beta (*P* value)	Model 10 beta (*P* value)	Percentage significant
Gender		−.272 (.001)	−.766 (<.001)		−.422 (.03)	−.365 (.005)						40
Age		.400 (<.001)	.234 (<.001)		.324 (<.001)	.130 (.009)	.322 (<.001)	.136 (<.001)	.138 (.012)	.285 (<.001)	.184 (<.001)	90
Education		.146 (<.001)			.315 (.001)	.195 (.001)	.082 (.002)	.095 (.01)			.139 (.008)	60
Occupation												
	Full-time student (reference)											
	Stay at home mom or dad				−.720 (.04)					−1.057 (<.001)		20
	Management	.546 (.004)							−1.675 (.002)			20
	Teacher or professor			−2.348 (.005)	−1.139 (.01)	.810 (.02)			−2.609 (.000)	−.949 (.02)		50
	Administrative, financial or clerical sales or service	.519 (.001)					.377 (.01)	.852 (.005)	-.894 (.035)			40
	Technologist or technical occupation	.532 (.003)										10
	Farming, forestry, fishing or mining	1.016 (<.001)			5.301 (<.001)		.400 (.04)			3.793 (<.001)		30
	Trades, transport or equipment operator				−1.564 (.02)	−1.047 (.007)	−.690 (<.001)				−.696 (.05)	40
	Processing, manufacturing or utilities					−.846 (.02)	−.641 (.001)					20
	Unemployed at present or on work leave	.479 (.008)							−.820 (.02)	−.587 (.02)		20
	Professional services (eg, certified accountant, lawyer, doctor)				−.856 (<.001)							10
	Occupation not listed	.703 (<.001)				.825 (.001)	.488 (.001)		−1.314 (.002)	−.945 (.001)	.647 (.004)	60

Regression Analysis: Indication-Specific Variables

In total, 10 indication-specific variables were tested for their association with posting behavior in the 2 addiction health interventions (Table 9).

Problem Drinking Intervention

In the problem drinking intervention, registrants had the option of selecting 1 of the 3 program goals. Compared with registrants who indicated that they wanted to cut down, *quit drinking* was positively and significantly associated with the number of posts in model 2 (beta=.463, *P*=.02). The option *not sure* was negatively and significantly associated with the number of posts in model 2 (beta=−. 460, *P*=.02) and model 7 (beta=−.509, *P*=.001).

Table 9. Statistically significant indication-specific independent variables (addiction interventions).

Independent Variables		Model 2 beta (*P* value)	Model 7 beta (*P* value)	Model 5 beta (*P* value)	Model 10 beta (*P* value)
Goal					
	Cut down (reference)			n/a[a]	n/a
	Quit drinking	.463 (.02)		n/a	n/a
	Not sure	−.460 (.02)	−0.509 (.001)	n/a	n/a
Smoking patterns					
	≥ one cigarette per day, occasional smoker, recently quit	n/a	n/a	.278 (.001)	
	Last cigarette: >24 hours, <24 hours	n/a	n/a	.534 (.002)	
	Cigarettes per day	n/a	n/a		
	Total years smoked	n/a	n/a	.040 (<.001)	.025 (.001)
	Minutes to first cigarette: >60, 31-60, 6-31, ≤5	n/a	n/a	.705 (<.001)	.625 (<.001)
	Past year quit attempts > 24 hours	n/a	n/a	−.048 (.02)	−.054 (.001)
	Number of cohabitant smokers	n/a	n/a		
	Fagerstrom dependency score (very low, low, moderate, high, very high)	n/a	n/a	0.657 (.001)	0.651 (<.001)

[a] n/a: not applicable.

Smoking Cessation Intervention

In model 5, increased cigarette consumption (smoking patterns) (beta=.278, *P*=.001) and having a cigarette within the past 24 hours (last cigarette) were positively and significantly associated with posting behavior (beta=.534, *P*=.002).

In both models, increases in total years smoked (beta=.040, *P*<.001; beta=.025, *P*=.001), decreases in minutes to first cigarette (beta=.705, *P*=.002; beta=.625, *P*<.001), and higher Fagerstrom dependency scores (beta=.657, *P*=.001; beta=.651, *P*<.001) were positively and significantly associated with posting behavior. Having a greater number of quit attempts was negatively and significantly associated with posting (beta = −.048, *P*=.02; −.054, *P*=.001).

Regression Analysis: Indication-Specific Variables in Two Mental Health Interventions

Ten indication-specific variables were tested for their association with posting behavior in the 2 mental health interventions. Whether a participant had *tried cognitive behavior therapy in the past* and was *using of the program with a health care* professional were asked in both mental health interventions (Table 10).

Past Cognitive Behavior Therapy Experience

In models 3, 4, and 9 posting behavior was positively and significantly associated with experience with CBT (beta= .851, *P*=.01; beta=1.118, *P*<.001; beta=.870, *P*<.001).

Table 10. Statistically significant indication-specific independent variables (mental health).

Independent variables	Model 3 beta (*P* value)	Model 8 beta (*P* value)	Model 4 beta (*P* value)	Model 9 beta (*P* value)
Depression rating past 2 weeks (0-10)			n/a[a]	n/a
Level of distress past 2 weeks (0-10)			n/a	n/a
Level of interference past 2 weeks (0-10)			n/a	n/a
Currently being treated			n/a	n/a
Tried cognitive behavior therapy in the past	.851 (.01)		1.118 (<.001)	.870 (<.001)
Number of attacks over past 2 weeks	n/a	n/a	.054 (.03)	
Using program with a health care professional				
Average fear rating during attack	n/a	n/a		−.099 (.01)
Attack interference in average daily life	n/a	n/a	.406 (<.001)	.224 (.01)
Attack causing avoidance	n/a	n/a		

[a] n/a: not applicable.

Depression Intervention

In the depression interventions, other than past CBT experience, there were no statistically significant associations with posting behavior.

Panic Disorder Intervention

In the panic disorder intervention, attacks interfering in average daily life were positively and significantly associated with posting behavior (beta=.406, *P*<.001; beta=.224, *P*=.01). In model 4, increases in number of attacks over the past 2 weeks were positively and significantly associated with posting (beta=.054, *P*=.03), and in model 9 average fear rating during an attack was negatively and significantly associated with posting (beta=−.099, *P*=.01).

Discussion

Principal Findings

Despite observable statistically significant results in demographic and indication-specific data, all regressions had low R^2 values, and their impact on superuser behavior was minimal. As mentioned previously, all models fail to explain the variance of the dependent variables.

Based on the results in 4 of the 10 models, females tend to post more than males. However, these results should be interpreted with caution as the impact was minimal (beta range=−.766 to −.272) and only statistically significant in all subject models. These results also do not confirm the gender of superusers.

Increased posting with age was positively and statistically significant in 9 of the 10 models, although the increase is negligible and should be interpreted with caution (beta range=.130-.400). For example, the analysis did not consider whether addiction treatment for smoking cessation, or if treatment for mental health issues, also coincides with age.

Although the impact is minimal, increased education was related to increases in posting behavior in 6 of the 10 models (beta range=.082-.315). The issue of education level and use of medical resources has a rich history in the literature and is nonconclusive. For example, one might assume that actors with higher levels should have better knowledge seeking skills and make limited use of DHSNs, or conversely, that actors with lower education levels and fewer formal resources would use DHSNs with greater intensity.

A recent qualitative review on factors affecting therapeutic compliance found the effect of education level to be equivocal [99]. While some studies found that patients with higher levels of education might have higher compliance, others found that patients with lower levels of education or no formal education were more compliant. The authors concluded that education level was not a good predictor of therapeutic compliance, and our findings reflect this in regards to education being associated with posting.

In the smoking cessation intervention, inexperienced quitters who have smoked longer, have increased dependency, and have recently quit, tend to post more. This supports past research indicating that the intervention’s DHSN primarily acts as a relapse prevention tool for new quitters [45,82]. If this finding is true it highlights the importance of detecting and supporting superusers as they primarily respond to, and support, new users.

It was interesting to note that *experience with cognitive behavior therapy* was associated with posting behavior in 3 of the 4 mental health models, though this impact was minimal (beta range=.851-1.118).

Future Research

The results of this study suggest that demographic or indication-specific variables have limited association with the creation of externalities in DHSNs. What, if anything, may be associated with posting behavior? If superusers are key to the growth and sustainability of DHSNs, how can they be detected?

The real-time assessment of phenotype, or observable traits resulting from the interaction of an individual in an environment, have recently been recognized as key to the next frontier of medicine [100]. Phenotypes differ from demographic and indication-specific data as they give insight on behavior. Although traditionally difficult to detect, some phenotypes are now being recognized through big data analysis.

For example, a recent study identified the ability to use natural language processing to detect phenotypes in electronic health records [101]. Another study found that an individual's personal attitudes including use of addictive substances, happiness, and sexual orientation can be detected through Facebook likes [102], and Instagram photos and Twitter feeds have been shown to contain predictive markers of depression [103,104].

DHSN content may contain rich sources of phenotypes as an post or an actor's profile may include avatars, images, badges or awards for participation, likes or other semiotic indicators of support from other members, or links to specific outside resources. Post content may be mined for specific keywords, phrases, or even tone. Time of post, time between posts, response to specific types of content or members, or other time-based interactions may also be indicative of specific behavior. Recent health care informatics research has also identified a relationship between increased systems use and outcomes, and a variety of unique system measures that may help categorize behaviors [105].

A challenge is that even if phenotypes can be predicted, risk-stratifying behavior may prove difficult. However, the medication adherence literature, which generally classifies patients as full compliers, partial compliers, or noncompliers may give insights on categorizing behavior similar to nonadherence [106] and research is beginning to investigate indication-specific factors that categorize patients and their motivations [107-110]. Future research into adherence to DHSNs might also consider the feasibility of stratifying actors according to real-time behavior.

In some respects, the low R^2 values in the models and lack of statistically significant variables in this study expose the limitations of big data. Popular belief holds that large data sets of survey data will contain insights and intelligence that have been previously unobtainable [111-113], and the promise of big data is so compelling that laymen are being encouraged to experiment with sophisticated techniques that previously required a high degree of training [114]. Whereas increased knowledge and interdisciplinary training and collaboration are certainly positive, as in this study, results from the analysis of large datasets pertaining to specific demographic characteristics or indication-specific variables may, at best, illustrate the complexity of predicting human behavior.

Strengths and Limitations

The results of this study are from "real world" social networks and the main strengths are the longevity of the DHSNs, the number of posts, the 4 separate indications, and that 2 of the social networks in the study were focused on mental health, and the remaining 2 on addictions.

Ideally, data from this study would be derived from a randomized controlled experiment. However, it would be difficult, if not impossible, to recruit a study population and execute a study in a similar sample. We are not aware of any other study in the health care literature with such an extensive and complete dataset, and as such, results should be interpreted accordingly.

A strength and limitation is that the populations analyzed are self-selecting populations that actively sought help. In the context of this study it was helpful to have datasets of active and engaged participants. However, these results may not be indicative of populations of patients in health plans, hospital networks, or mass public health campaigns.

A limitation to this study is that demographic and indication-specific data was self-report. Self-report data is common in digital health studies, and the consensus is that data from subjects is at least as reliable as pencil-and-paper questionnaires [115-122]. However, due to the anonymous nature and nonrandomization of study subjects, results should be interpreted with caution.

Conclusions

Based on the large number of study participants, variation in DHSN theme, and extensive time-period, we did not find strong evidence that demographic characteristics or indication severity sufficiently explain the variability in number of posts per actor. Researchers should investigate alternative methods and models that may identify individuals who promote DHSN growth.

Conflicts of Interest

Trevor van Mierlo is the CEO & Founder of Evolution Health Systems. Evolution Health owns and manages digital health interventions, including the applications analyzed in this study.

Multimedia Appendix 1

[PPTX File, 2MB - jmir_v19i2e40_app1.pptx]

References

1. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. Br Med J 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
2. Jadad AR, Enkin MW, Glouberman S, Groff P, Stern A. Are virtual communities good for our health? Br Med J 2006 Apr 22;332(7547):925-926 [FREE Full text] [doi: 10.1136/bmj.332.7547.925] [Medline: 16627491]
3. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]
4. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The role of social network technologies in online health promotion: a narrative review of theoretical and empirical factors influencing intervention effectiveness. J Med Internet Res 2015;17(6):e141 [FREE Full text] [doi: 10.2196/jmir.3662] [Medline: 26068087]
5. Pew Research Center. Washington, DC; 2016 Nov 11. Social Media Fact Sheet URL: http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/ [accessed 2017-01-23] [WebCite Cache ID 6nn23udaa]
6. Cole J, Watkins C, Kleine D. Health advice from Internet discussion forums: how bad is dangerous? J Med Internet Res 2016 Jan 06;18(1):e4 [FREE Full text] [doi: 10.2196/jmir.5051] [Medline: 26740148]
7. Myneni S, Cobb N, Cohen T. In pursuit of theoretical ground in behavior change support systems: analysis of peer-to-peer communication in a health-related online community. J Med Internet Res 2016 Feb 02;18(2):e28. [doi: 10.2196/jmir.4671] [Medline: 26839162]
8. Robinson TN, Walters PA. Health-net: an interactive computer network for campus health promotion. J Am Coll Health 1986 Jun;34(6):284-285. [doi: 10.1080/07448481.1986.9938952] [Medline: 3745693]
9. Bender JL, Jimenez-Marroquin MC, Ferris LE, Katz J, Jadad AR. Online communities for breast cancer survivors: a review and analysis of their characteristics and levels of use. Support Care Cancer 2013 May;21(5):1253-1263. [doi: 10.1007/s00520-012-1655-9] [Medline: 23179491]
10. Brennan PF, Moore SM, Smyth KA. The effects of a special computer network on caregivers of persons with Alzheimer's disease. Nurs Res 1995;44(3):166-172. [Medline: 7761293]
11. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop P. Online social networks and smoking cessation: a scientific research agenda. J Med Internet Res 2011;13(4):e119 [FREE Full text] [doi: 10.2196/jmir.1911] [Medline: 22182518]
12. Ploderer B, Smith W, Howard S, Pearce J, Borland R. Patterns of support in an online community for smoking cessation. In: C&T '13. 2013 Presented at: Proceedings of the 6th International Conference on Communities and Technologies; June 29 - July 02, 2013; Munich, Germany p. 26-35.
13. Takahashi Y, Uchida C, Miyaki K, Sakai M, Shimbo T, Nakayama T. Potential benefits and harms of a peer support social network service on the internet for people with depressive tendencies: qualitative content analysis and social network analysis. J Med Internet Res 2009;11(3):e29 [FREE Full text] [doi: 10.2196/jmir.1142] [Medline: 19632979]
14. Wright K. Social support within an on-line cancer community: an assessment of emotional support, perceptions of advantages and disadvantages, and motives for using the community from a communication perspective. J Appl Commun Res 2002 Jan;30(3):195-209. [doi: 10.1080/00909880216586]
15. Wicks P, Keininger DL, Massagli MP, de la Loge C, Brownstein C, Isojärvi J, et al. Perceived benefits of sharing health data between people with epilepsy on an online platform. Epilepsy Behav 2012 Jan;23(1):16-23 [FREE Full text] [doi: 10.1016/j.yebeh.2011.09.026] [Medline: 22099528]
16. Conrad P, Bandini J, Vasquez A. Illness and the Internet: from private to public experience. Health (London) 2016 Jan;20(1):22-32. [doi: 10.1177/1363459315611941] [Medline: 26525400]
17. Patton M. Forbes. 2015 Jun 29. Health Care Costs Rise Faster Than Inflation URL: http://www.forbes.com/sites/mikepatton/2015/06/29/u-s-health-care-costs-rise-faster-than-inflation/ [accessed 2017-01-25] [WebCite Cache ID 6nn3ENIDh]
18. Woolhandler S, Campbell T, Himmelstein DU. Costs of health care administration in the United States and Canada. N Engl J Med 2003 Aug 21;349(8):768-775. [doi: 10.1056/NEJMsa022033] [Medline: 12930930]
19. Schneider EL. The Aging of America. J Am Med Assoc 1990 May 02;263(17):2335. [doi: 10.1001/jama.1990.03440170057036]
20. Zweifel P, Felder S, Meiers M. Ageing of population and health care expenditure: a red herring? Health Econ 1999 Sep;8(6):485-496. [Medline: 10544314]
21. Getzen TE. Population aging and the growth of health expenditures. J Gerontol 1992 May;47(3):S98-104. [Medline: 1573213]
22. de Meijer C, Wouterse B, Polder J, Koopmanschap M. The effect of population aging on health expenditure growth: a critical review. Eur J Ageing 2013 May 15;10(4):353-361. [doi: 10.1007/s10433-013-0280-x]
23. Chisholm-Burns MA, Spivey CA. The 'cost' of medication nonadherence: consequences we cannot afford to accept. J Am Pharm Assoc (2003) 2012;52(6):823-826. [doi: 10.1331/JAPhA.2012.11088] [Medline: 23229971]
24. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [Medline: 15076819]

25. Viswanathan M, Golin CE, Jones CD, Ashok M, Blalock SJ, Wines RC, et al. Interventions to improve adherence to self-administered medications for chronic diseases in the United States: a systematic review. Ann Intern Med 2012 Dec 4;157(11):785-795. [doi: 10.7326/0003-4819-157-11-201212040-00538] [Medline: 22964778]
26. Forissier T, Firlik K. Capgemini. 2012. Estimated Annual Pharmaceutical Revenue Loss Due to Medication Non-Adherence URL: https://www.capgemini.com/resource-file-access/resource/pdf/Estimated_Annual_Pharmaceutical_Revenue_Loss_Due_to_Medication_Non-Adherence.pdf [accessed 2017-01-25] [WebCite Cache ID 6nn41PJoS]
27. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769 [FREE Full text] [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
28. Glasgow RE, McKay HG, Piette JD, Reynolds KD. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? Patient Educ Couns 2001 Aug;44(2):119-127. [Medline: 11479052]
29. Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006 May 16;144(10):742-752. [Medline: 16702590]
30. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
31. Harrison MI, Koppel R, Bar-Lev S. Unintended consequences of information technologies in health care--an interactive sociotechnical analysis. J Am Med Inform Assoc 2007;14(5):542-549 [FREE Full text] [doi: 10.1197/jamia.M2384] [Medline: 17600093]
32. Emirbayer M, Goodwin J. Network analysis, culture, and the problem of agency. AJS 1994;99(6):1411-1454.
33. Grabher G, Stark D. Organizing diversity: evolutionary theory, network analysis and postsocialism. Reg Stud 1997 Jul;31(5):533-544. [doi: 10.1080/00343409750132315]
34. Leibenstein H. Bandwagon, snob, and veblen effects in the theory of consumers' demand. Q J Econ 1950;64(2):183-207.
35. Luke DA, Harris JK. Network analysis in public health: history, methods, and applications. Annu Rev Public Health 2007;28:69-93. [doi: 10.1146/annurev.publhealth.28.021406.144132] [Medline: 17222078]
36. Safranek R. The Emerging Role of Social Media in Political and Regime Change. In: ProQuest Discover Guides. Ann Arbor: ProQuest; 2012:1-14.
37. Uzzi B. The sources and consequences of embeddedness for the economic performance of organizations: the network effect. Am Sociol Rev 1996;61(4):674-698.
38. Newman M. The structure and function of complex networks. SIAM Rev 2003 Jan;45(2):167-256. [doi: 10.1137/S003614450342480]
39. Robins G, Pattison P, Kalish Y, Lusher D. An introduction to exponential random graph (p*) models for social networks. Soc Networks 2007 May;29(2):173-191. [doi: 10.1016/j.socnet.2006.08.002]
40. Snijders TA, van de Bunt GG, Steglich CE. Introduction to stochastic actor-based models for network dynamics. Soc Networks 2010 Jan;32(1):44-60. [doi: 10.1016/j.socnet.2009.02.004]
41. Albert R, Barabási A. Statistical mechanics of complex networks. Rev Mod Phys 2002 Jan 30;74(1):47-97. [doi: 10.1103/RevModPhys.74.47]
42. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: a comparison of the 90-9-1 Principle and Zipf's Law. Internet Interv 2014 Oct;1(4):165-168. [doi: 10.1016/j.invent.2014.09.003]
43. Kossinets G, Watts DJ. Empirical analysis of an evolving social network. Science 2006 Jan 6;311(5757):88-90 [FREE Full text] [doi: 10.1126/science.1116869] [Medline: 16400149]
44. van MT, Hyatt D, Ching AT. Mapping power law distributions in digital health social networks: methods, interpretations, and practical implications. J Med Internet Res 2015;17(6):e160 [FREE Full text] [doi: 10.2196/jmir.4297] [Medline: 26111790]
45. van Mierlo T, Hyatt D, Ching AT. Employing the Gini coefficient to measure participation inequality in treatment-focused Digital Health Social Networks. Netw Model Anal Health Inform Bioinform 2016;5(1):32 [FREE Full text] [doi: 10.1007/s13721-016-0140-7] [Medline: 27840788]
46. Su W. Integrating and Mining Virtual Communities Across Multiple Online Social Networks: Concepts, Approaches and Challenges. In: IEEE Computer Society Press. 2014 Presented at: The Fourth International Conference on Digital Information and Communication Technology and its Applications; May 6, 2014; Bangkok p. 199-204. [doi: 10.1109/dictap.2014.6821682]
47. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From help-seekers to influential users: a systematic review of participation styles in online health communities. J Med Internet Res 2015 Dec 01;17(12):e271 [FREE Full text] [doi: 10.2196/jmir.4705] [Medline: 26627369]
48. Cunningham JA, van Mierlo T, Fournier R. An online support group for problem drinkers: AlcoholHelpCenter.net. Patient Educ Couns 2008 Feb;70(2):193-198. [doi: 10.1016/j.pec.2007.10.003] [Medline: 18022340]
49. Jones R, Sharkey S, Smithson J, Ford T, Emmens T, Hewis E, et al. Using metrics to describe the participative stances of members within discussion forums. J Med Internet Res 2011;13(1):e3 [FREE Full text] [doi: 10.2196/jmir.1591] [Medline: 21239373]

50. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian cancer society's smokers' helpline online and StopSmokingCenter.net. J Med Internet Res 2012;14(3):e66 [FREE Full text] [doi: 10.2196/jmir.1854] [Medline: 22732103]
51. 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. [doi: 10.1136/amiajnl-2013-002282] [Medline: 24449805]
52. Baye MR. Advanced Topics in Business Strategy. In: Managerial EconomicsBusiness Strategy. 7 ed. New York, NY: The McGraw-Hill Companies Inc; 2010.
53. Katz ML, Shapiro C. Network externalities, competition, and compatibility. Am Econ Rev 1985;75(3):424-440.
54. Lin K, Lu H. Why people use social networking sites: an empirical study integrating network externalities and motivation theory. Comput Human Behav 2011 May;27(3):1152-1161. [doi: 10.1016/j.chb.2010.12.009]
55. Andriani P, McKelvey B. Beyond Gaussian averages: redirecting international business and management research toward extreme events and power laws. J Int Bus Stud 2007 Oct 18;38(7):1212-1230. [doi: 10.1057/palgrave.jibs.8400324]
56. Newman M. Power laws, Pareto distributions and Zipf's law. Contemp Phys 2005 Sep;46(5):323-351. [doi: 10.1080/00107510500052444]
57. Gabaix X. Power laws in Economics and Finance. Annu Rev Econ 2009 Sep 02;1(1):255-294. [doi: 10.1146/annurev.economics.050708.142940]
58. Nielsen J. NNGroup. Fremont, CA: Nielson Normal Group; 2009 Oct 09. The 90-9-1 Rule for Participation Inequality: Lurkers vs. Contributors in Internet Communities URL: https://www.nngroup.com/articles/participation-inequality/ [accessed 2017-01-25] [WebCite Cache ID 6nn8X2Sms]
59. Arthur C. TheGuardian.: Guardian News and Media Ltd; 2006. What is the 1% rule? URL: https://www.theguardian.com/technology/2006/jul/20/guardianweeklytechnologysection2 [accessed 2017-01-25] [WebCite Cache ID 6nn8kUFF7]
60. Sanders R. The Pareto principle: its use and abuse. Jnl of Product & Brand Mgt 1992 Jun;1(2):37-40. [doi: 10.1108/10610429210036762]
61. Buntain C, Golbeck J. Identifying social roles in reddit using network structure. In: WWW '14 Companion. 2014 Presented at: The 23rd International Conference on World Wide Web; April 07 - 11, 2014; Seoul, Korea p. 615-620. [doi: 10.1145/2567948.2579231]
62. Graham T, Wright S. Discursive equality and everyday talk online: the impact of “Superparticipants”. J Comput-Mediat Comm 2013 May 20;19(3):625-642. [doi: 10.1111/jcc4.12016]
63. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966]
64. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Cunningham JA, Griffiths KM. Community structure of a mental health internet support group: modularity in user thread participation. JMIR Ment Health 2016;3(2):e20 [FREE Full text] [doi: 10.2196/mental.4961]
65. Alcoholhelpcenter. Toronto: Evolution Health Systems The Alcohol Help Center URL: http://www.alcoholhelpcenter.net/ [accessed 2017-01-24] [WebCite Cache ID 6nn9xKov6]
66. DepressionCenter. Toronto: Evolution Health Systems The Depression Center URL: http://www.DepressionCenter.net/ [accessed 2017-01-22] [WebCite Cache ID 6nnA1woVo]
67. PanicCenter. Toronto: Evolution Health Systems The Panic Center URL: http://www.PanicCenter.net/ [accessed 2017-01-24] [WebCite Cache ID 6nnA9xaVa]
68. StopSmokingCenter. Toronto: Evolution Health The Stop Smoking Center URL: http://www.StopSmokingCenter.net/ [accessed 2017-01-24] [WebCite Cache ID 6nnAEfz5S]
69. Cunningham J. Internet Evidence-Based Treatments. In: Miller PE, editor. Evidence-Based Addiction Treatment. Burlington: Academic Press; 2009:379-397.
70. Cunningham JA. Comparison of two internet-based interventions for problem drinkers: randomized controlled trial. J Med Internet Res 2012 Aug 01;14(4):e107 [FREE Full text] [doi: 10.2196/jmir.2090] [Medline: 22954459]
71. Cunningham JA, Humphreys K, Kypri K, van Mierlo T. Formative evaluation and three-month follow-up of an online personalized assessment feedback intervention for problem drinkers. J Med Internet Res 2006 Apr 12;8(2):e5 [FREE Full text] [doi: 10.2196/jmir.8.2.e5] [Medline: 16867968]
72. Cunningham JA, Selby P, van Mierlo T. Integrated online services for smokers and drinkers? Use of the check your drinking assessment screener by participants of the Stop Smoking Center. Nicotine Tob Res 2006 Dec;8 Suppl 1:S21-S25. [Medline: 17491167]
73. Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
74. Cunningham JA, Wild TC, Cordingley J, van Mierlo MT, Humphreys K. Twelve-month follow-up results from a randomized controlled trial of a brief personalized feedback intervention for problem drinkers. Alcohol Alcohol 2010;45(3):258-262 [FREE Full text] [doi: 10.1093/alcalc/agq009] [Medline: 20150170]
75. Davis J. The Panic Center. Child Adolesc Ment Health 2007 Feb;12(1):49-50. [doi: 10.1111/j.1475-3588.2007.00437_3.x]

76. Doumas DM, McKinley LL, Book P. Evaluation of two Web-based alcohol interventions for mandated college students. J Subst Abuse Treat 2009 Jan;36(1):65-74. [doi: 10.1016/j.jsat.2008.05.009] [Medline: 18657941]
77. 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 Mar 26;7(1):e7 [FREE Full text] [doi: 10.2196/jmir.7.1.e7] [Medline: 15829479]
78. Farvolden P, McBride C, Bagby RM, Ravitz P. A Web-based screening instrument for depression and anxiety disorders in primary care. J Med Internet Res 2003;5(3):e23 [FREE Full text] [doi: 10.2196/jmir.5.3.e23] [Medline: 14517114]
79. McDonnell DD, Lee H, Kazinets G, Moskowitz JM. Online recruitment of targeted populations: lessons learned from a smoking cessation study among Korean Americans. Soc Mar Q 2010 Sep 26;16(3):2-22. [doi: 10.1080/15245004.2010.500441]
80. Pike KJ, Rabius V, McAlister A, Geiger A. American Cancer Society's QuitLink: randomized trial of Internet assistance. Nicotine Tob Res 2007 Mar;9(3):415-420. [doi: 10.1080/14622200701188877] [Medline: 17365773]
81. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. J Med Internet Res 2008 Nov 21;10(5):e45 [FREE Full text] [doi: 10.2196/jmir.1008] [Medline: 19033146]
82. Selby P, van Mierlo T, Voci SC, Parent D, Cunningham JA. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. J Med Internet Res 2010 Aug 18;12(3):e34 [FREE Full text] [doi: 10.2196/jmir.1340] [Medline: 20719739]
83. van Mierlo T, Fournier R, Jean-Charles A, Hovington J, Ethier I, Selby P. I'll txt U if I have a problem: how the Société Canadienne du cancer in Quebec applied behavior-change theory, data mining and agile software development to help young adults quit smoking. PLoS One 2014;9(3):e91832 [FREE Full text] [doi: 10.1371/journal.pone.0091832] [Medline: 24647098]
84. O'Donnell A, Wallace P, Kaner E. From efficacy to effectiveness and beyond: what next for brief interventions in primary care? Front Psychiatry 2014;5:113 [FREE Full text] [doi: 10.3389/fpsyt.2014.00113] [Medline: 25221524]
85. Herbert J, Forman E. The Evolution of Cognitive Behavioral Therapy: the rise of psychological acceptancemindfullness. In: Acceptance and Mindfulness in Cognitive Behavior Therapy: Understanding and applying the new therapies. New York: Wiley; 2011:1-26.
86. Deterding S. Gamification. Interactions 2012 Jul 01;19(4):14-17. [doi: 10.1145/2212877.2212883]
87. Erickson P, Butters J, Walko K, Butterill D, Reggie C, Fischer B, et al. CAMH. Toronto; 2002. CAMH and Harm Reduction: A Background Paper on its Meaning and Application for Substance Use Issues URL: http://www.camh.ca/en/hospital/about_camh/influencing_public_policy/public_policy_submissions/harm_reduction/Pages/harmreductionbackground.aspx [accessed 2017-01-25] [WebCite Cache ID 6nnGOCNMc]
88. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [Medline: 6392204]
89. Miller W, Rollnick S. Motivational Interviewing: Preparing People to Change. 2nd ed. New York: The Guilford Press; 2002.
90. Moreira MT, Smith LA, Foxcroft D. Social norms interventions to reduce alcohol misuse in university or college students. Cochrane Database Syst Rev 2009 Jul 08(3):CD006748. [doi: 10.1002/14651858.CD006748.pub2] [Medline: 19588402]
91. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
92. Sanchez-Craig M. Drink Wise: How to quit drinking or cut down. A self-help book. Toronto: Addiction Research Foundation; 1995.
93. Kreuter MW, Skinner CS. Tailoring: what's in a name? Health Educ Res 2000 Feb;15(1):1-4. [Medline: 10788196]
94. Prochaska JO, DiClemente CC. Transtheoretical therapy: toward a more integrative model of change. Psychotherapy Theory Research & Practice 1982;19(3):276-288. [doi: 10.1037/h0088437]
95. EUR-Lex. 1995. Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data URL: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31995L0046 [accessed 2017-01-25] [WebCite Cache ID 6nnHTzPRC]
96. Laws-lois.justice.: Government of Canada; 2015 Jun 23. Personal Information Protection and Electronic Documents Act URL: http://laws-lois.justice.gc.ca/eng/acts/P-8.6/ [accessed 2017-01-25] [WebCite Cache ID 6nnHfHDt8]
97. HHS.: Office for Civil Rights Headquarters; 2003 May. Summary of the HIPAA Privacy Rule URL: https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html?language=es [accessed 2017-01-25] [WebCite Cache ID 6nnI7DhCF]
98. World Medical Association Declaration. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Med Assoc 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
99. Jin J, Sklar GE, Min Sen Oh V, Chuen LS. Factors affecting therapeutic compliance: a review from the patient's perspective. Ther Clin Risk Manag 2008 Feb;4(1):269-286 [FREE Full text] [Medline: 18728716]
100. Ausiello D, Lipnick S. Real-time assessment of wellness and disease in daily life. Big Data 2015 Sep 01;3(3):203-208 [FREE Full text] [doi: 10.1089/big.2015.0016] [Medline: 26487991]

101. Yu S, Liao KP, Shaw SY, Gainer VS, Churchill SE, Szolovits P, et al. Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. J Am Med Inform Assoc 2015 Sep;22(5):993-1000 [FREE Full text] [doi: 10.1093/jamia/ocv034] [Medline: 25929596]
102. Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci 2013 Mar 11;110(15):5802-5805 [FREE Full text] [doi: 10.1073/pnas.1218772110]
103. Reece A, Danforth C. Cornell University Library. 2016. Instagram photos reveal predictive markers of depression URL: https://arxiv.org/abs/1608.03282 [accessed 2017-01-25] [WebCite Cache ID 6nnJNclEp]
104. Reece A, Reagan A, Lix K, Dodds P, Danforth C, Langer E. Cornell University Library. 2016 Aug 27. Forecasting the onset and course of mental illness with Twitter data URL: https://arxiv.org/abs/1608.07740 [accessed 2017-01-25] [WebCite Cache ID 6nnJU0JZy]
105. Su W, Chih M. Is More eHealth System Use Better for Cancer PatientsFamily Caregivers? A Literature Review. 2016 Presented at: American Medical Informatics Association 2016 Annual Symposium; November 12-16, 2016; Chicago, IL. [doi: 10.13140/RG.2.1.4077.7200]
106. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clinical Therapeutics 2001 Aug;23(8):1296-1310. [doi: 10.1016/s0149-2918(01)80109-0]
107. Cramer JA. A systematic review of adherence with medications for diabetes. Diabetes Care 2004 May;27(5):1218-1224. [Medline: 15111553]
108. van Mierlo T, Fournier R, Ingham M. Targeting medication non-adherence behavior in selected autoimmune diseases: a systematic approach to digital health program development. PLoS One 2015;10(6):e0129364 [FREE Full text] [doi: 10.1371/journal.pone.0129364] [Medline: 26107637]
109. Hill J. Adherence with drug therapy in the rheumatic diseases Part two: measuring and improving adherence. Musculoskeletal Care 2005;3(3):143-156. [doi: 10.1002/msc.4] [Medline: 17042003]
110. Julian LJ, Yelin E, Yazdany J, Panopalis P, Trupin L, Criswell LA, et al. Depression, medication adherence, and service utilization in systemic lupus erythematosus. Arthritis Rheum 2009 Feb 15;61(2):240-246 [FREE Full text] [doi: 10.1002/art.24236] [Medline: 19177526]
111. boyd D, Crawford K. Critical questions for big data. Inf Commun Soc 2012;15(5):662-679 [FREE Full text] [doi: 10.1080/1369118X.2012.678878]
112. Poynter R. VisionCritical.: Vision Critical Communications Inc; 2013. Big Data successes and limitations: What researchers and marketers need to know URL: https://www.visioncritical.com/big-data-successes-and-limitations/ [accessed 2017-01-25] [WebCite Cache ID 6nnKetKuc]
113. Wu X, Zhu X, Wu GQ, Ding W. Data mining with big data. IEEE Trans Knowl Data Eng 2014 Jan;26(1):97-107. [doi: 10.1109/TKDE.2013.109]
114. Luo G. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. Netw Model Anal Health Inform Bioinforma 2016;5(1):1-16. [doi: 10.1007/s13721-016-0125-6]
115. Carlbring P, Brunt S, Bohman S, Austin D, Richards J, Öst L, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. Comput Human Behav 2007 May;23(3):1421-1434. [doi: 10.1016/j.chb.2005.05.002]
116. Fanning J, McAuley E. A comparison of tablet computer and paper-based questionnaires in healthy aging research. JMIR Res Protoc 2014;3(3):e38 [FREE Full text] [doi: 10.2196/resprot.3291] [Medline: 25048799]
117. Hohwü L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective. J Med Internet Res 2013 Aug 26;15(8):e173 [FREE Full text] [doi: 10.2196/jmir.2595] [Medline: 23978658]
118. Riva G, Teruzzi T, Anolli L. The use of the internet in psychological research: comparison of online and offline questionnaires. Cyberpsychol Behav 2003 Feb;6(1):73-80. [doi: 10.1089/109493103321167983] [Medline: 12650565]
119. Smith AB, King M, Butow P, Olver I. A comparison of data quality and practicality of online versus postal questionnaires in a sample of testicular cancer survivors. Psychooncology 2013 Jan;22(1):233-237. [doi: 10.1002/pon.2052] [Medline: 21954102]
120. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. J Med Internet Res 2011 Dec 04;13(4):e109 [FREE Full text] [doi: 10.2196/jmir.1593] [Medline: 22155721]
121. Wu RC, Thorpe K, Ross H, Micevski V, Marquez C, Straus SE. Comparing administration of questionnaires via the internet to pen-and-paper in patients with heart failure: randomized controlled trial. J Med Internet Res 2009 Feb 06;11(1):e3 [FREE Full text] [doi: 10.2196/jmir.1106] [Medline: 19275979]
122. Brigham J, Lessov-Schlaggar CN, Javitz HS, Krasnow RE, McElroy M, Swan GE. Test-retest reliability of web-based retrospective self-report of tobacco exposure and risk. J Med Internet Res 2009 Aug 11;11(3):e35 [FREE Full text] [doi: 10.2196/jmir.1248] [Medline: 19674962]

Abbreviations

DHSN: Digital health social network

Edited by G Eysenbach; submitted 07.07.16; peer-reviewed by B Carron-Arthur, A Graham, WC Su; comments to author 18.11.16; revised version received 31.12.16; accepted 13.01.17; published 17.02.17.

Please cite as:
van Mierlo T, Li X, Hyatt D, Ching AT
Demographic and Indication-Specific Characteristics Have Limited Association With Social Network Engagement: Evidence From 24,954 Members of Four Health Care Support Groups
J Med Internet Res 2017;19(2):e40
URL: http://www.jmir.org/2017/2/e40/
doi:10.2196/jmir.6330
PMID:28213340

Viewpoint

Navigating Ethics in the Digital Age: Introducing Connected and Open Research Ethics (CORE), a Tool for Researchers and Institutional Review Boards

John Torous[1,2], MD; Camille Nebeker[3,4,5], MS, EdD

[1]Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States
[2]Division of Clinical Informatics, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States
[3]Department of Family Medicine and Public Health, Division of Behavioral Medicine, UC San Diego School of Medicine, La Jolla, CA, United States
[4]Center for Wireless and Population Health Systems, Qualcomm Institute, UC San Diego, La Jolla, CA, United States
[5]Scripps Translational Science Institute, La Jolla, CA, United States

Corresponding Author:
Camille Nebeker, MS, EdD
Department of Family Medicine and Public Health
Division of Behavioral Medicine
UC San Diego School of Medicine
9500 Gilman Drive, MC-0725
La Jolla, CA, 92093-0725
United States
Phone: 1 858 534 7786
Fax: 1 858 534 4642
Email: nebeker@eng.ucsd.edu

Abstract

Research studies that leverage emerging technologies, such as passive sensing devices and mobile apps, have demonstrated encouraging potential with respect to favorably influencing the human condition. As a result, the nascent fields of mHealth and digital medicine have gained traction over the past decade as demonstrated in the United States by increased federal funding for research that cuts across a broad spectrum of health conditions. The existence of mHealth and digital medicine also introduced new ethical and regulatory challenges that both institutional review boards (IRBs) and researchers are struggling to navigate. In response, the Connected and Open Research Ethics (CORE) initiative was launched. The CORE initiative has employed a participatory research approach, whereby researchers and IRB affiliates are involved in identifying the priorities and functionality of a shared resource. The overarching goal of CORE is to develop dynamic and relevant ethical practices to guide mHealth and digital medicine research. In this Viewpoint paper, we describe the CORE initiative and call for readers to join the CORE Network and contribute to the bigger conversation on ethics in the digital age.

(J Med Internet Res 2017;19(2):e38) doi:10.2196/jmir.6793

KEYWORDS

ethics, research; mHealth; research; mobile technologies; telemedicine

Introduction

The transformative potential of mobile, digital, and passive sensing technologies to observe or intervene with various health domains is now well established. As the field transitions from discussing to now testing this potential, the role of clinical studies and evidence generation will assume predominance.

While there are many new challenges in conducting technology-enhanced research, even navigating study approval via the institutional review board (IRB) ethical and regulatory evaluation process remains a primary challenge for many. Whether the existing IRB system is able to support the ethical review and subsequent advancement of research using these new and emerging technologies was recently questioned [1]. To address ethical challenges in this new age of technology-enhanced research, we outline a potential solution.

Health-related research using tools, such as social networks, mobile phones apps, and wearable passive sensors, offer the potential to collect unprecedented amounts of real-time data outside of the clinic and in real-world or “free-living” environments. From monitoring location using a mobile phone's global positioning sensors to inferring socialness based on call and text message logs, mobile and passive sensing technologies offer a new window into human behavior. Yet such broad data capture can also have numerous unintended consequences [2,3]. Research participants may be unaware of the nature, scope, and granularity of data collected and what information they are actually consenting to provide. Wearable recording may capture the sounds and images of nearby people not involved in the research, or "bystanders," due to their close proximity to a research participant [4]. Text message logs may also reveal personal communications from and with bystanders, introducing potential privacy concerns and legal challenges [5]. It is easy to imagine many more examples where both personal and community privacy boundaries can easily be infringed upon and compromised.

Many researchers are eager to study the clinical potential of mobile and digital tools, but they may not be aware of all the potential risks to participants, or of means to mitigate them [6,7]. IRBs have thus been placed in the difficult role of evaluating these research proposals and ensuring they are both safe and ethical. Yet, like many researchers, IRB members themselves are not experts in this emerging field and may struggle to evaluate the safety and ethical dimensions of submissions that often involve novel uses of various technologies. In many cases, there are actually no standards, best practices, or demonstrated safety mechanisms to guide either researchers or IRB risk assessment or management strategies. Researchers may thus feel uncomfortable with explaining digital or mHealth benefits, risks, and management strategies in the IRB protocol application and, likewise, IRBs may feel uncomfortable approving study plans. The outcome is a lengthy IRB review process that may result in either the underprotection or overprotection of research participants [3,4].

Given the evolving state of knowledge and growing interest in using technologies, one simple solution is sharing of developing practices and examples of successful digital- and mHealth-related IRB submissions. The idea of sharing resources to guide the ethical conduct of research using novel strategies is not new. In fact, sharing guidance was initiated years ago when researchers started to use the Internet to support scientific inquiry. For example, the Association of Internet Research is a membership organization that aims to foster ethical and professional Internet research by offering guidance and education to academics, including faculty and students [8]. Similarly, the British Psychological Society has developed ethical guidance for Internet-mediated research [9], as has INVOLVE [10], which is a Web-based resource launched in the United Kingdom 20 years ago with support from the National Institute for Health Research. While these resources focus primarily on Internet research, there are other organizations thinking about ethical dimensions of nanoengineering, robotics, and artificial intelligence, yet not necessarily focusing on research with human participants. Several years ago, we recognized that researchers and ethics review boards (eg, IRBs or research ethics boards) may benefit from having access to a community of stakeholders with expertise in various digital and mHealth tools. The initial goal was to create an accessible and dynamic resource to bridge a growing gap between researchers who capture personal health data via mobile (apps), imaging (eg, Microsoft's SenseCam wearable camera), pervasive sensing (eg, wearable, ingestible, and environmental technology), social media (eg, Twitter, Instagram, and Facebook), and geolocation tracking (eg, global positioning systems and geographic information systems) tools and the IRBs charged with reviewing these studies.

Introducing the Connected and Open Research Ethics Initiative

The Connected and Open Research Ethics (CORE) initiative was launched in 2015 with support from the Robert Wood Johnson Foundation (Princeton, NJ). The CORE is a free, Web-based resource that aims to convene stakeholders in the digital-mHealth ecosystem to collectively shape dynamic and responsive ethical and responsible research practices. Using a participatory approach to inform the CORE design and function, the CORE team invited input from individuals representing interdisciplinary, cross-disciplinary, and cross-sector perspectives with a vested interest in advancing dynamic and responsive ethical standards. Focus groups and key informant interviews were convened with IRB representatives and researcher stakeholders to inform the initial CORE Platform functionality and design. The CORE Platform, released in 2016, hosts a growing global network of over 200 individuals representing 10 countries and a majority of the United States with expertise in privacy, technology, bioethics, research ethics, regulations, sciences, engineering, and even a few participants. The key features include a Forum where Network members can share informational resources and post or answer questions, and a Resource Library where researchers can upload language used in their IRB-approved protocol application and informed consent documents. The goal is to help other researchers who are beginning to use new digital tools in research and who want to see examples of successful IRB protocol and consent language, and receive feedback from experts when writing their own IRB applications.

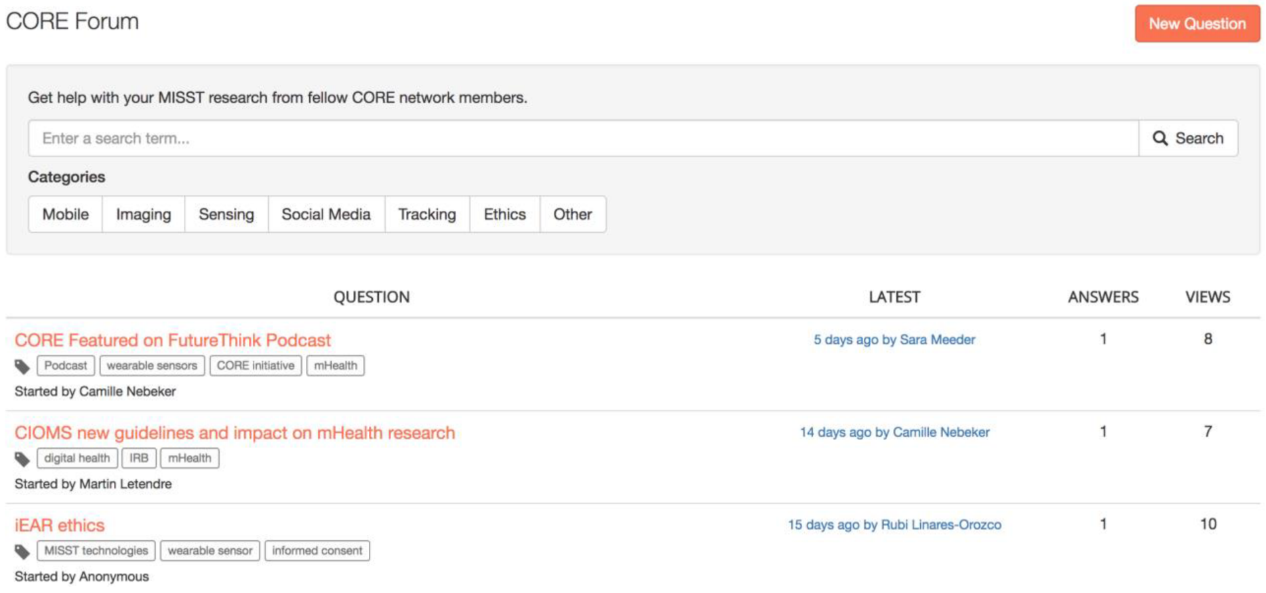
Likewise, IRBs that are beginning to review mHealth and digital medicine research studies can post questions on the CORE Forum (Figure 1), as well as contribute to or search the Resource Library (Figure 2), to see what others have found to be acceptable. This saves time and, ideally, increases the consistency for how IRBs evaluate and mitigate potential risks to research participants.

That being said, we cannot be certain that an IRB approval means that the ethical review is beyond reproach. As the CORE community begins to share resources, we expect stakeholders to chime in when a potential risk has been overlooked or, likewise, when the approved protocol appears overly conservative. The CORE is also where new resources (eg, institutional policies or guidelines) can be shared and ideas explored, which may lead to potential collaborations.

As with any innovation, early adopters to the CORE also serve as beta testers who help to improve utility and functionality. As we enter this new frontier where vast and granular amounts of personal health data are collected in real time and 24/7, we look to the CORE Network members to begin shaping how to do this research in a manner that is informed, ethical, and responsive to participants' interests. The CORE Resource Library is designed so that as tools or methods become obsolete, they will drop off and the new innovations will percolate to the top. Moreover, the CORE has a rating system so that if a Network member notes a gap in the data security plan, they can chime in and inform the community of a better practice.

We invite readers of the *Journal of Medical Internet Research* to join the CORE Network to share knowledge, access resources, and contribute to shaping the ethics for 21st century research. Get started by visiting CORE's website [11] and sign in to create an account. Once you are on the CORE Platform, browse the Resource Library and visit the Forum to engage in discussion with others in the CORE community.

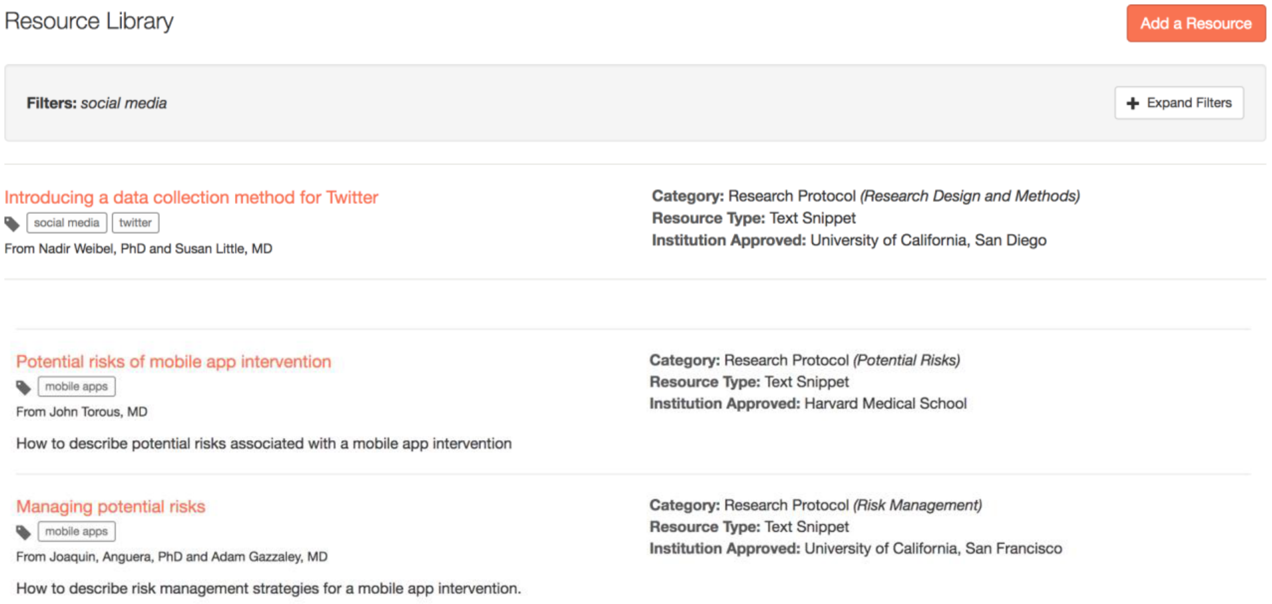
Figure 1. Connected and Open Research Ethics (CORE) Forum screenshot. MISST: mobile imaging, pervasive sensing, social media and location tracking.



Get help with your MISST research from fellow CORE network members.

QUESTION	LATEST	ANSWERS	VIEWS
CORE Featured on FutureThink Podcast Podcast, wearable sensors, CORE initiative, mHealth Started by Camille Nebeker	5 days ago by Sara Meeder	1	8
CIOMS new guidelines and impact on mHealth research digital health, IRB, mHealth Started by Martin Letendre	14 days ago by Camille Nebeker	1	7
iEAR ethics MISST technologies, wearable sensor, informed consent Started by Anonymous	15 days ago by Rubi Linares-Orozco	1	10

Figure 2. Connected and Open Research Ethics (CORE) Resource Library screenshot.



Filters: *social media*

Introducing a data collection method for Twitter
social media, twitter
From Nadir Weibel, PhD and Susan Little, MD

Category: Research Protocol *(Research Design and Methods)*
Resource Type: Text Snippet
Institution Approved: University of California, San Diego

Potential risks of mobile app intervention
mobile apps
From John Torous, MD
How to describe potential risks associated with a mobile app intervention

Category: Research Protocol *(Potential Risks)*
Resource Type: Text Snippet
Institution Approved: Harvard Medical School

Managing potential risks
mobile apps
From Joaquin, Anguera, PhD and Adam Gazzaley, MD
How to describe risk management strategies for a mobile app intervention.

Category: Research Protocol *(Risk Management)*
Resource Type: Text Snippet
Institution Approved: University of California, San Francisco

Acknowledgments

This work relates to a grant from the Robert Wood Johnson Foundation to develop the “Connected and Open Research Ethics [CORE] (principal investigator: CN, #72876, 2015-2017) and the UC San Diego Chancellor’s Interdisciplinary Collaboratory Fellowship Program.”

Conflicts of Interest

CN is the principal investigator of the Connected and Open Research Ethics project. JT is an unpaid member of the CORE Advisory Committee.

References

1. Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, et al. Reimagining human research protections for 21st century science. J Med Internet Res 2016 Dec 22;18(12):e329 [FREE Full text] [doi: 10.2196/jmir.6634] [Medline: 28007687]
2. Torous J, Keshavan M, Gutheil T. Promise and perils of digital psychiatry. Asian J Psychiatr 2014 Aug;10:120-122. [doi: 10.1016/j.ajp.2014.06.006] [Medline: 25042968]
3. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
4. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med 2016 Dec;6(4):577-586. [doi: 10.1007/s13142-016-0426-4] [Medline: 27688250]
5. Nebeker C, Linares-Orozco R, Crist K. A multi-case study of research using mobile imaging, sensing and tracking technologies to objectively measure behaviorthical issues and insights to guide responsible research practice. J Res Admin 2015;46(1):118-137.
6. Pisani AR, Wyman PA, Mohr DC, Perrino T, Gallo C, Villamar J, et al. Human subjects protection and technology in prevention science: selected opportunities and challenges. Prev Sci 2016 Aug;17(6):765-778 [FREE Full text] [doi: 10.1007/s11121-016-0664-1] [Medline: 27220838]
7. Torous J, Roberts LW. The ethical use of mobile health technology in clinical psychiatry. J Nerv Ment Dis 2017 Jan;205(1):4-8. [doi: 10.1097/NMD.0000000000000596] [Medline: 28005647]
8. Association of Internet Researchers. Chicago, IL: AoIR; 2017. URL: http://aoir.org/ [accessed 2017-02-01] [WebCite Cache ID 6nxQJtYK8]
9. Hewson C, Buchanan T, Brown I, Coulson N, Hagger-Johnson G, Joinson A, et al. Ethics guidelines for Internet-mediated research. Leicester, UK: The British Psychological Society; 2013. URL: http://www.bps.org.uk/system/files/Public%20files/inf206-guidelines-for-internet-mediated-research.pdf [accessed 2016-12-26] [WebCite Cache ID 6n3hyX1p7]
10. National Institute for Health Research. INVOLVE. Southampton, UK: INVOLVE; 2015. URL: http://www.invo.org.uk/ [accessed 2017-02-06] [WebCite Cache ID 6o4p1YTI7]
11. CORE Platform. Lafayette, CA: The Regents of the University of California; 2016. URL: https://thecore-platform.ucsd.edu/ [accessed 2017-02-01] [WebCite Cache ID 6nxRw1vdm]

Abbreviations

CORE: Connected and Open Research Ethics
IRB: institutional review board

Edited by G Eysenbach; submitted 13.10.16; peer-reviewed by M Larsen, N Berry, TR Soron; comments to author 18.12.16; revised version received 28.12.16; accepted 03.01.17; published 08.02.17.

Please cite as:
Torous J, Nebeker C
Navigating Ethics in the Digital Age: Introducing Connected and Open Research Ethics (CORE), a Tool for Researchers and Institutional Review Boards
J Med Internet Res 2017;19(2):e38
URL: http://www.jmir.org/2017/2/e38/
doi:10.2196/jmir.6793
PMID:28179216

Original Paper

Screening for Child Sexual Exploitation in Online Sexual Health Services: An Exploratory Study of Expert Views

Victoria Spencer-Hughes[1], MSc, FFPH; Jonathan Syred[1], MSc; Alison Allison[2], MSc; Gillian Holdsworth[3], PGDE, MSc, MPH, MBChB, FFPH; Paula Baraitser[1,4], MD, FFSRH, FFPH

[1]King's College London, London, United Kingdom
[2]Independent Researcher, London, United Kingdom
[3]Lambeth and Southwark Public Health Directorate, London, United Kingdom
[4]King's College Hospital NHS Foundation Trust, London, United Kingdom

Corresponding Author:
Paula Baraitser, MD, FFSRH, FFPH
King's College London
Weston Education Centre
10 Cutcombe Road
London, SE5 9RJ
United Kingdom
Phone: 44 2078485168
Fax: 44 2078485769
Email: paula.baraitser@kcl.ac.uk

Abstract

Background: Sexual health services routinely screen for child sexual exploitation (CSE). Although sexual health services are increasingly provided online, there has been no research on the translation of the safeguarding function to online services. We studied expert practitioner views on safeguarding in this context.

Objective: The aim was to document expert practitioner views on safeguarding in the context of an online sexual health service.

Methods: We conducted semistructured interviews with lead professionals purposively sampled from local, regional, or national organizations with a direct influence over CSE protocols, child protection policies, and sexual health services. Interviews were analyzed by three researchers using a matrix-based analytic method.

Results: Our respondents described two different approaches to safeguarding. The “information-providing” approach considers that young people experiencing CSE will ask for help when they are ready from someone they trust. The primary function of the service is to provide information, provoke reflection, generate trust, and respond reliably to disclosure. The approach values online services as an anonymous space to test out disclosure without commitment. The “information-gathering” approach considers that young people may withhold information about exploitation. Therefore, services should seek out information to assess risk and initiate disclosure. This approach values face-to-face opportunities for individualized questioning and immediate referral.

Conclusions: The information-providing approach is associated with confidential telephone support lines and the information-gathering approach with clinical services. The approach adopted online will depend on ethos and the range of services provided. Effective transition from online to clinic services after disclosure is an essential element of this process and further research is needed to understand and support this transition.

(J Med Internet Res 2017;19(2):e30) doi:10.2196/jmir.5911

KEYWORDS

Internet; child abuse, sexual; adolescent health services; sexually transmitted diseases; risk assessment

Introduction

Safeguarding children is “the action we take to promote the welfare of children and to protect them from harm” [1]. One form of harm is child sexual exploitation (CSE), which involves those younger than 18 years in exploitative situations, contexts, and relationships in which they receive something (eg, gifts or money) for engaging in sexual activity [2]. Determining the incidence of CSE is complex and underreporting is common

[3], but it is estimated that there were 16,500 children at risk in England during the period from April 2010 to March 2011 [4].

Young people experiencing sexual exploitation may use sexual health services even when they have disengaged with other statutory services [2]. Therefore, sexual health services have an important role in identifying CSE. They fulfill this role through routine history taking based on national guidelines for users younger than age 18 and clearly specified referral pathways when concerns are identified [2]. In England, the age of sexual consent is 16 years and all those younger than 18 years are considered at risk of CSE.

Sexual health services are increasingly provided online [5]. The online interface is particularly attractive to young people who value the accessibility, convenience, and discretion of online services [6,7]. Online sexual health services come in many forms from those that only provide sexually transmitted infection tests to comprehensive provision of testing, treatment, and contraception in association with text, telephone, and webchat support and referral to clinical services. The common element of these services is the lack of face-to-face contact, but they may offer different levels of clinical support and links to other relevant services. The lack of face-to-face contact raises specific concerns about processes for identification of risk of CSE. Some of the signs of risk for CSE are difficult to assess online, such as poor self-care, injuries, emotional symptoms, trauma symptoms, or self-harming behavior [4]. At present, online services are limited to those older than 16 years.

There is little national or international guidance on safeguarding within an online sexual health service. A literature search using the terms (or variations of) safeguard, online, Internet, and Web on the databases Allied and Complementary Medicine Database (AMED), British Nursing Index (BNI), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, Health Business Elite, Health Management Information Consortium (HMIC), MEDLINE, and PsychINFO found no directly relevant information to guide policy development.

Due to the lack of published evidence to inform policy development in this area, we interviewed local and national expert practitioners in safeguarding and sexual health to document their views on safeguarding young people using online sexual health services.

Methods

Ethical Approval

Ethical approval was received from King’s College London Ethics Committee (REC Reference Number *:* BDM/13/14-102).

Study Design

Semistructured interviews were completed with expert practitioner stakeholders. Following national guidance on involving the potential beneficiaries of research in research design, data collection, and analysis [8], we worked with young people from a local government-funded school as cointerviewers. Participation was voluntary and open to all students aged 16 to 18 years living in the local area who wanted to be involved in health services research that affected young people. The cointerviewers developed four questions for the interview schedule through a collaborative process led by two of the researchers (JS and AA). One young person attended each of the interviews, where possible, to ask these questions.

Sampling and Recruitment

A purposive sample of local and national practitioner experts on safeguarding, young people’s rights, and sexual health services was recruited. We identified possible respondents using a snowballing technique in which local experts identified relevant organizations and then purposively sampled from this group for maximum variability to include a wide range of organizations that had direct influence over CSE protocols and child protection policies (Table 1). All statutory and National Health Service organizations in the geographical area were included. Participants chose to give their views either on behalf of their named organization or in a personal capacity.

Table 1. Organizations and professional roles of interviewees

Organization	Role	Regional/National
Brook: national sexual health charity for young people	Chief executive	National
English National Chlamydia Screening Programme	Director	National
English National Chlamydia Screening Programme	Quality assurance manager	National
Children's charity offering support and protection to children being abused	Message board manager speaking in personal capacity	National
Local statutory organizations with responsibility for safeguarding children	Chair	Regional
Local statutory organizations with responsibility for safeguarding children	Development manager	Regional
Local statutory organizations with responsibility for safeguarding children	Director of children's social care	Regional
Local statutory organizations with responsibility for safeguarding children	Public health consultant	Regional
Local statutory organizations with responsibility for safeguarding children	Senior manager service development	Regional
Hospital services provider	Sexual health promotion manager	Local
Community sexual health services provider	Safeguarding lead clinician	Local
Sexual assault referral center	Senior clinician	Regional
Regional health care commissioning body	Safeguarding nurse	Regional
Youth/housing services	Safeguarding manager	Local

Data Collection

A total of 14 interviews, lasting 40 to 60 minutes, were completed at the researcher's or the participants' place of work with a young person researcher present in eight of them. The interviews were recorded, transcribed, and then analyzed in Nvivo10 (NVivo qualitative data analysis Software; QSR International Pty Ltd, version 10, 2012).

The interview included four sections: experience of safeguarding in clinic and online services, differences between the two safeguarding environments, and the scenario presented in Textbox 1.

Textbox 1. Scenario for interview.

> Apple is 16 years old and has been having sex. She does not feel able to go to her family GP or a local clinic to get an STI test but would like to check that she is healthy. Apple goes online to the online sexual health service website, completes the online risk assessment [which asks questions about the same issues as clinics do].

The discussion was repeated with Apple aged 15 years and when she had ticked a box to say that her partner was in a position of trust (eg, a teacher or youth worker) or that alcohol had affected her sexual behavior. These were chosen as factors that would trigger further questioning in a clinic environment as part of standard protocols.

A final section included questions from the young person interviewer and an opportunity for the participant to add anything else they felt relevant.

Data Analysis

Interviews were analyzed using the framework approach [9], a matrix-based analytic method that classifies and organizes data according to key themes, subthemes, and emergent categories [10]. The interview texts were read and reread by PB and VSH before developing coding categories which were then refined through three rounds of coding and modification through discussion between PB and VSH (see Multimedia Appendix 1 for coding categories). A third researcher (JS) checked the coding of the first six transcripts (43%, 6/14) to identify initial inconsistencies in coding. As the coding categories became increasingly robust and familiar to the coders, this checking function was no longer required. Differences were resolved through discussion. The cointerviewers commented on a preliminary analysis and draft conclusions.

Results

No respondent described young people visiting sexual health services primarily to disclose concerns about sexual exploitation. It was assumed that information about sexual exploitation would be identified during an interaction primarily focused on sexually transmitted infection or pregnancy.

Six themes emerged from the data that were grouped into two sections:

Issues relevant to the identification of CSE. The themes in this section were access, trust, and strategies for data collection. We identified two strategies that we termed “information gathering” and “information providing.”

Issues relevant to the response to CSE. The themes in this section were the consistency and effectiveness (in terms of stopping the exploitation) of responses to disclosure and online services as part of a wider system.

Issues Relevant to the Identification of Child Sexual Exploitation

Respondents described increased access to online services with no risk of being seen using the clinic, no need to travel to the service, adapt to clinic opening hours, or wait to be seen. However, online services require private Internet access, health literacy, proficiency in English, and confidential access to mail at home.

Trust in the service was described as essential to support disclosure of CSE. The development of trust may require several visits, including testing visits. Trust includes confidence that services are nonjudgmental, will keep personal information confidential, and will reliably respond to the self-identified needs of young people at the pace they choose. In clinics, skilled clinicians may build trust quickly within the consultation and have the advantage of verbal and nonverbal cues to communicate their reliability. However, questioning can be overwhelming for young people and health professionals might lack time or skills to question appropriately. Online services provide a less pressured environment than clinics for building trust. They provide consistency with standardized questions and responses, time for young people to reflect, and a potentially anonymous space for young people to test service responses. However, online services remain faceless and it is easier to ignore key information or disengage with the process:

> *We know that some of the reasons that children phone Childline are because they don’t want to go in through a door. We know that some of the reasons that people contact Brook first without anybody going to see them are that they want to test that you’re okay and then they will build that trust.* [Respondent 6]

If very good and clear information was put up about safeguarding, it would give young people an opportunity to really read it and think about it and be able to make decisions on that...it’s sometimes quite confrontational where you’ve got a face-to-face contact in a clinic and you’re being told and it’s a bit scary anyway, and you’re not taking it on board. (Respondent 13)

During analysis of the theme of trust two differing approaches to safeguarding emerged: information providing and information gathering (see Textbox 2).

Textbox 2. Information-providing versus information-gathering approaches to identification of child sexual exploitation.

Information-Gathering Approach

Identify contextual and nonverbal clues

"If a young person is delaying their responses or being fidgety or looks scared or concerned when a topic of conversation is raised, I think that gives professionals who have years of experience dealing with young people quite a lot of information" (Respondent 4)

Collect information from a range of sources

"They might not be registered with a GP...But then they might be known to lots of other agencies so that (the online service) could be a good way of tracking them down" (Respondent 13)

Professional risk assessment

"Then you just really have to cue into pauses and stumbling over words. And inconsistencies as well...But sometimes it doesn't all add up...you have to be very direct in saying, 'First of all you said this and now you're saying that'" Respondent 8

Young people will withhold information to prevent referral

"Sometimes they won't give you that detail because they know what's going to happen with that information, so they'd rather not" (Respondent 9)

People lie

"Ultimately you're fighting against people who may be intentionally deceptive" (Respondent 4)

Information-Giving Approach

Creating a safe space supports disclosure:

"If they're...not harried and harassed, if they're given the opportunity to actually go and access something where they feel somebody will listen to them kindly" (Respondent 13)

Use the information given

"So, if they don't reveal it, they don't reveal it. I've always said in safeguarding, you can only do what you do on the information you're given and actually if you start digging around people might just go, poof and not be seen ever again" (Respondent 13)

Young people assess their own risk

"Where people are risk assessing themselves you're relying on people to be seeking the level of advice that they need" (Respondent 6)

Young people will seek help when they need it

"If a person feels that they can trust this entity providing these services they may disclose what they are ready to disclose, if they are not ready to disclose then they will not give that information" (Respondent 11)

People will tell the (broad) truth

"People...will generally tell you the truth within limits, adults who go to the doctor who asks how much they drink will often say more than they're supposed to but not quite as much as they do, by doing that enough to know a doctor's going to say, 'You need to worry about this a little bit'" (Respondent 6)

The information-providing approach considers that young people experiencing CSE will ask for help from someone they trust. The role of a service is to meet the criteria for young people to have sufficient trust to seek help when they need it, prompt recognition of risk by providing information, and to respond appropriately and in a timely fashion. The information-providing approach values the online service as an option that provides a safe space to test out disclosure without commitment.

The information-gathering approach considers that young people may withhold information out of concern that it might trigger an intervention or judgmental response. Therefore, services are obliged to seek out information and to form a judgment about the young person's situation that is independent of the young person's account. The information-gathering approach values the extended opportunities for individualized questioning to assess risk and challenge perceptions of sexual relationships and the extended contact afforded by presence in the clinic.

Issues Relevant to the Response to Child Sexual Exploitation

When CSE is identified, a rapid response that stops the exploitation is an important outcome. Respondents noted that face-to-face services are better placed to provide immediate protection than online services. These views reference the advantages of the young person's physical presence in the clinic at the time of disclosure that facilitates immediate involvement of social services and or the police, whereas contact online is more fragile especially when there is ambivalence about disclosure. This is particularly the case if inaccurate contact details are provided online. A dichotomy emerged in this analysis in which participants described the advantages of a user-controlled response to CSE or a provider-controlled one. Those who favored the provider-controlled response referenced the need to respond immediately:

> *If somebody has come out with that you have got to grab hold of it while it is there because you might not get another chance.* [Respondent 14]

Those who favored a user-controlled response referenced the futility of a response without user support because it was likely to lead to the retraction of the disclosure.

Participants felt strongly that to offer effective responses to the disclosure of CSE online services must be integrated into a wider clinical system. Standalone online services may offer limited knowledge of, or connections with, local support services; therefore, there are reduced opportunities for effective referral or follow-up in face-to-face services.

Discussion

Child sexual exploitation is a subset of child sexual abuse (CSA). The nature and dynamics of CSA make it extremely difficult for young people to disclose exploitation [11]. Barriers to disclosure include dependency, strategies employed by perpetrators to maintain silence, feelings of guilt and responsibility, and fear of not being believed [11,12]. Although most CSA is first disclosed to peers and family members, approximately 10% is disclosed to professionals, including sexual health professionals [13]. Disclosure of CSE is further complicated by the young person not recognizing they are being exploited. An environment that encourages recognition of exploitation and disclosure providing a prompt response is essential to support the young person to stop the exploitation and to reduce the risk of long-term negative outcomes [11].

Online services may offer advantages in supporting the disclosure of CSE by offering consistent information, signaling an appropriate response to disclosure; facilitating initial and repeat (often testing) visits; and by providing time for reflection. Online services provide immediate, consistent, and nonjudgmental responses with a sense of safety and control that comes through remote access [14] and the possibility of “space for negotiation” rather than immediate response [15]. More sensitive information is reported via computer questionnaires than face-to-face interviews [16,17], and self-completed questionnaires can be effective in identifying CSE [18]. The “faceless” and “voiceless” nature of these services are important for young people who wish to discuss personal problems [14], particularly when these problems are stigmatizing [19]. Confidential telephone counseling services that require no identifying information are important resources for young people in crisis situations [20], such as those contemplating suicide [21].

Our distinction between the information-providing and the information-gathering approaches can help further thinking about this issue. Figure 1 depicts the spectrum of remote support services for young people who are experiencing CSE. Organizations on the left of the diagram are less likely to have contact details for the young person and focus predominately on providing support until the young person is ready to disclose. Organizations on the right of the diagram are more likely to have contact details, be in a position to crosscheck information, and to insist on referrals.

Our research suggests that depending on their approach and the range of services they provide (anonymous telephone support, online chat, online clinical services that store more or less information), online services may position themselves more to the left or the right of the diagram in Figure 1 and, therefore, may adopt more of an information-gathering or an information-providing approach. The approach should be clearly signaled to users so that they are aware of the consequences of disclosure.

However, at some point, the response to disclosure of CSE will require face-to-face contact to involve the relevant services (eg, social workers and the police) and to stop the exploitation. This requires transfer of the conversation from the online service to a different context. This is the key concern about safeguarding online. We have found no published evidence about referrals between online sexual health services and clinics, although referrals from telephone helplines on all topics are successful in approximately 50% of cases [20].

From our interviews with expert practitioners, we can see a consensus that services embedded within wider clinical or organizational structures may be more effective at supporting safeguarding of children than standalone services.

Further research is needed on how to affect this transfer. This research could usefully focus on any of the following questions:

What proportion of those who disclose CSE online can be effectively referred to face-to-face services?

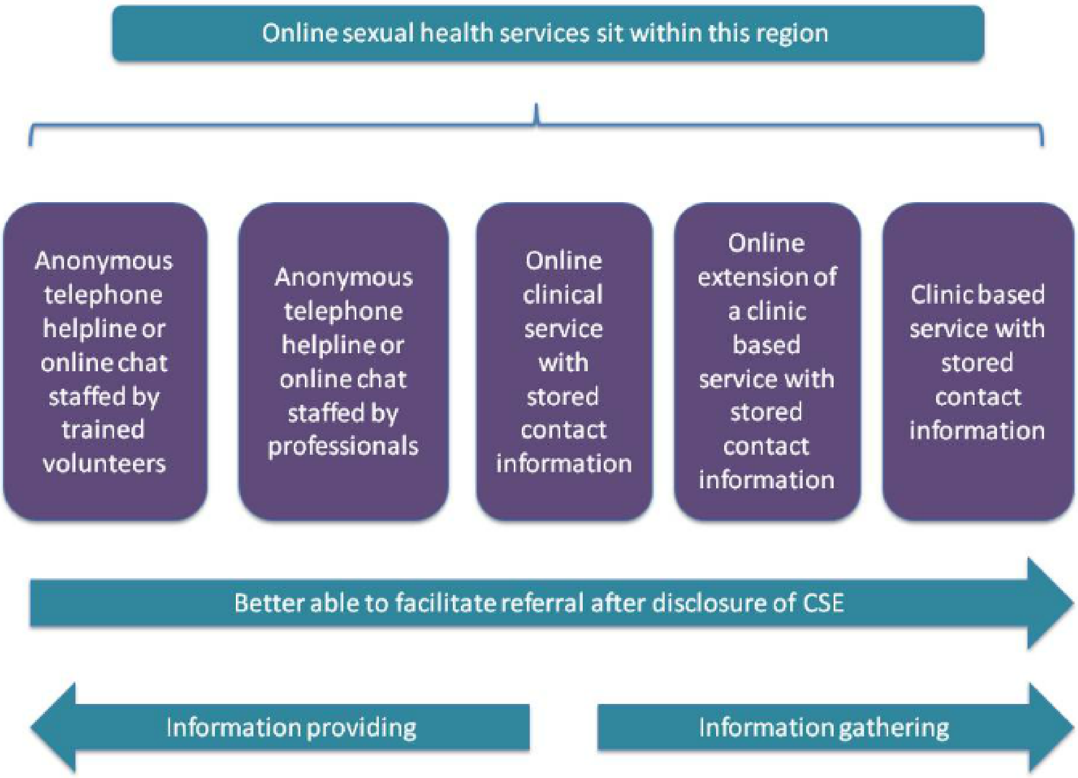
Is an information-gathering or an information-providing approach more effective in (1) supporting disclosure of CSE and (2) referral to face-to-face services?

Are online sexual health services more effective in supporting disclosure of CSE (whatever the approach adopted) than face-to-face services?

None of the participants had experience working in an online sexual health service, although the respondents from Brook and the Children’s charity had extensive experience providing support via telephone and online chat to young people in need of help. Participants come from a range of professional backgrounds linked to safeguarding. Some of the respondents worked almost exclusively with young people who are at risk of, or are, being sexually exploited, whereas others worked with young people seeking contraceptive or sexual health services. Young people were not included as participants in this study, although they were involved in later service development work.

In conclusion, some elements of online sexual health services may facilitate disclosure of CSE. Effective transition from online to clinic services after disclosure is an essential element of this process and further research is needed to understand how this transition can be supported.

Figure 1. Spectrum of approaches to supporting disclosure of child sexual exploitation.



Acknowledgments

We would like to thank all those interviewed for this paper for giving their time. We would like to thank Lindsay Pain for support in coding interview transcripts.

We would like to thank the students and staff of St Martins in the Field School, Lambeth, for their involvement in the study. We would like to extend our sincere thanks to Debbie Timehin, Dubheasa Lanipekun, Mariam Adeniji, Miriam Bamba, Natalie Adoteye, Nathanial Arias Fontanil, Rachel Ogunbayo, and Zoe Wynne for their committed work as cointerviewers and their contribution to this study.

This study was funded by Guy’s and St Thomas’ charity as part of the development and evaluation of an online sexual health service for Lambeth and Southwark.

Conflicts of Interest

PB and GH are directors of SH:24, a community-interest company providing online sexual health services.

Multimedia Appendix 1

Coding categories.

[PDF File (Adobe PDF File), 14KB - jmir_v19i2e30_app1.pdf]

References

1. HM Government. Working Together to Safeguard Children: A Guide to Inter-Agency Working to Safeguard and Promote the Welfare of Children. 2015. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/419595/Working_Together_to_Safeguard_Children.pdf [WebCite Cache ID 6gszT4KIO]
2. Rogstad K, Johnston G. Spotting the Signs: A National Proforma for Identifying Risk of Child Sexual Exploitation in Sexual Health Services. Cheshire, UK: BASSH/Brook; 2014. URL: https://www.bashh.org/documents/Spotting-the-signs-A%20national%20proforma%20Apr2014.pdf [accessed 2017-01-19] [WebCite Cache ID 6ndQLyjYV]

3. Academy of Medical Royal Colleges. Child Sexual Exploitation: Improving Recognition and Response in Health Settings. 2014 Sep. URL: http://www.aomrc.org.uk/publications/reports-a-guidance [accessed 2016-04-19] [WebCite Cache ID 6gt071wTN]
4. Berelowitz S, Firmin C, Edwards G, Gulyurtlu G. "I Thought I Was the Only One. The Only One in the World" The Office of the Children's Commissioner's Inquiry into Child Sexual Exploitation in Gangs and Groups: Interim Report. London: The Office of the Children's Commissioner; 2012 Nov. URL: http://deraioe.ac.uk/16067/1/FINAL_REPORT_FOR_WEBSITE_Child_Sexual_Exploitation_in_Gangs_and_Groups_Inquiry_Interim_Report__21_11_12.pdf [accessed 2016-04-19] [WebCite Cache ID 6gt0KfsqK]
5. Woodhall S, Sile B, Talebi A, Nardone A, Baraitser P. Internet testing for Chlamydia trachomatis in England, 2006 to 2010. BMC Public Health 2012 Dec 19;12:1095 [FREE Full text] [doi: 10.1186/1471-2458-12-1095] [Medline: 23253518]
6. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006 Jun 23;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
7. Lorimer K, McDaid L. Young men's views toward the barriers and facilitators of Internet-based Chlamydia trachomatis screening: qualitative study. J Med Internet Res 2013 Dec 03;15(12):e265 [FREE Full text] [doi: 10.2196/jmir.2628] [Medline: 24300158]
8. Department of Health. Research Governance Framework for Health and Social Care. 2005. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/139565/dh_4122427.pdf, [WebCite Cache ID 6gt0eRkZg]
9. Ritchie J. Analysing qualitative data. In: Ritchie S, editor. The Qualitative Researcher's Companion. London: Routledge; 1994:173-194.
10. Ritchie J, Lewis J. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage Publications; 2003.
11. Paine ML, Hansen DJ. Factors influencing children to self-disclose sexual abuse. Clin Psychol Rev 2002 Mar;22(2):271-295. [Medline: 11806022]
12. McElvaney R, Greene S, Hogan D. To tell or not to tell? factors influencing young people's informal disclosures of child sexual abuse. J Interpers Violence 2014 Mar;29(5):928-947. [doi: 10.1177/0886260513506281] [Medline: 24288188]
13. Allnock D. Children and Young People Disclosing Sexual Abuse: An Introduction to the Research. Research Briefing. London: NSPCC; 2010. URL: http://www.childmatters.org.nz/file/Diploma-Readings/Block-2/Sexual-Abuse/3.4-children-and-young-people-disclosing-sexual-abuse-updated.pdf [accessed 2016-04-19] [WebCite Cache ID 6gt0sEd5P]
14. Fukkink R, Hermanns JM. Children's experiences with chat support and telephone support. J Child Psychol Psychiatry 2009 Jun;50(6):759-766. [doi: 10.1111/j.1469-7610.2008.02024.x] [Medline: 19207634]
15. Vincent S, Daniel B. An analysis of children and young people's calls to ChildLine about abuse and neglect: a study for the Scottish Child Protection Review. Child Abuse Rev 2004 May 7;13(2):158-171. [doi: 10.1002/car.840]
16. Dolezal C, Marhefka S, Santamaria E, Leu C, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. Arch Sex Behav 2012 Apr;41(2):401-410 [FREE Full text] [doi: 10.1007/s10508-011-9769-6] [Medline: 21604065]
17. Millstein S, Irwin C. Acceptability of computer-acquired sexual histories in adolescent girls. J Pediatr 1983 Nov;103(5):815-819. [Medline: 6631616]
18. Milne L, Collin-Vézina D. Disclosure of sexual abuse among youth in residential treatment care: a multiple informant comparison. J Child Sex Abus 2014;23(4):398-417. [doi: 10.1080/10538712.2014.896841] [Medline: 24640965]
19. Caplan S, Turner J. Bringing theory to research on computer-mediated comforting communication. Comput Hum Behav 2007 Mar;23(2):985-998. [doi: 10.1016/j.chb.2005.08.003]
20. Stein D, Lambert MJ. Telephone counseling and crisis intervention: a review. Am J Community Psychol 1984 Feb;12(1):101-126. [Medline: 6711489]
21. King R, Nurcombe B, Bickman L, Hides L, Reid W. Telephone counselling for adolescent suicide prevention: changes in suicidality and mental state from beginning to end of a counselling session. Suicide Life Threat Behav 2003;33(4):400-411. [Medline: 14695055]

Abbreviations

CSA: child sexual abuse
CSE: child sexual exploitation

Edited by G Eysenbach; submitted 04.05.16; peer-reviewed by D Allknock, J Reid; comments to author 31.07.16; revised version received 03.10.16; accepted 29.10.16; published 14.02.17.

Please cite as:
Spencer-Hughes V, Syred J, Allison A, Holdsworth G, Baraitser P
Screening for Child Sexual Exploitation in Online Sexual Health Services: An Exploratory Study of Expert Views
J Med Internet Res 2017;19(2):e30
URL: http://www.jmir.org/2017/2/e30/
doi:10.2196/jmir.5911
PMID:28196790

Viewpoint

The Effect of the General Data Protection Regulation on Medical Research

John Mark Michael Rumbold[1*], GDL, MB ChB, PhD; Barbara Pierscionek[2,3*], PhD, LLM

[1]Faculty of Science, Engineering and Computing, Kingston University London, Kingston upon Thames, United Kingdom
[2]School of Science and Technology, 50 Shakespeare Street, Nottingham Trent University, Nottingham NG1 4FQ, United Kingdom
[3]Faculty of Science, Engineering and Computing, Penrhyn Road, Kingston University London, Kingston upon Thames KT1 2EE, United Kingdom
[*]all authors contributed equally

Corresponding Author:
John Mark Michael Rumbold, GDL, MB ChB, PhD
Faculty of Science, Engineering and Computing
Kingston University London
Penrhyn Road
Kingston upon Thames,
United Kingdom
Phone: 44 020 841 74918
Fax: 44 208 417 9000
Email: J.Rumbold@Kingston.ac.uk

Abstract

Background: The enactment of the General Data Protection Regulation (GDPR) will impact on European data science. Particular concerns relating to consent requirements that would severely restrict medical data research have been raised.

Objective: Our objective is to explain the changes in data protection laws that apply to medical research and to discuss their potential impact.

Methods: Analysis of ethicolegal requirements imposed by the GDPR.

Results: The GDPR makes the classification of pseudonymised data as personal data clearer, although it has not been entirely resolved. Biomedical research on personal data where consent has not been obtained must be of substantial public interest.

Conclusions: The GDPR introduces protections for data subjects that aim for consistency across the EU. The proposed changes will make little impact on biomedical data research.

(J Med Internet Res 2017;19(2):e47) doi:10.2196/jmir.7108

KEYWORDS

pseudonymity; anonymity; untraceability; privacy-preserving protocols; informatics; data reporting; data protection; research ethics

Overview

There have been significant developments in European Union (EU) data protection law recently that will have an impact on health care professionals, particularly those engaged in research and audit. The General Data Protection Regulation (GDPR) has replaced the current legislation and comes into full effect in 2018 [1]. The implications for the handling of health care data of the GDPR will be discussed in this paper. Despite the recent referendum vote in the United Kingdom to leave the EU, the GDPR will continue to be relevant to the United Kingdom, whether this is due to cooperation in European projects or because the United Kingdom continues to be a member of the European Economic Area (EEA).

The Data Protection Directive

Currently the relevant law in the United Kingdom is the Data Protection Act 1998, which is the United Kingdom's transposition of the Data Protection Directive (DPD). European directives are not directly enforceable, requiring member states to pass legislation to comply with their requirements. There are derogations (legal exemptions) for research, which in the case of the United Kingdom have been criticized for being too broad. The Local and Regional Development Planning Kantor report for the European Commission criticizes the United Kingdom for disregard of the limitations, stating that the Data Protection Act blatantly violates the Directive by adding “medical research” to the list of medical purposes [2]. The DPD requires a “substantial public interest” for member states to add to the derogations for processing of sensitive personal data (Article 8.4).

Differences between EU member states can result in research ethics committees in United Kingdom denying permission for National Health Service (NHS) data to be transferred to other EU countries (the opposite might also be the case in some circumstances) [3]. These differences have also contributed to the passage of the GDPR as part of the Digital Single Market strategy [4].

The Law as It Will Be From 2018: The General Data Protection Regulation

The text of the GDPR has recently been agreed after a prolonged trilogue between the European Commission, Parliament, and the Council of Ministers [5]. This legislation will replace the national transpositions of the DPD. Regulations are directly enforceable across the EU. The GDPR comes into full effect on May 25, 2018, although member states are permitted minor differences in interpretation (the European Court of Justice is the ultimate arbiter). This legislation has the potential to affect projects using research data banks and Big Data [6,7]. There had been concerns that a clause inserted by the European Parliament requiring specific consent would prevent significant long-term epidemiological research taking place in the future [8], but this was rejected and the agreed text permits broad consent to “certain areas of research when in keeping with recognized ethical standards” (Recital 33) [9]. Broad consent is not blanket or open consent [10] although some commentators argue that blanket or open consent is acceptable for biobank and databank research as the risks are minimal and do not vary for different projects [11]. Another possibility is consent to a form of governance [12]. Open consent without any ongoing regulation or communication about proposed projects would be potentially problematic. Dynamic consent offers advantages for an engaged community of participants but might not be considered beneficial by some individuals [13].

The derogations for research without consent have been expanded to specifically include medical research where “in the public interest” (Recital 51). How public interest will be defined has not been elaborated, but European jurisprudence demands member states satisfy a high threshold where human rights are involved (eg, a “pressing social need” [14]). This standard would not be required for the conduct of medical research using databanks, but it might exclude all commercial research for “me too” drug development (drugs that offer no advantages over drugs already on the market), arrangements that have no evidence of benefit sharing, or simply require that projects address issues of public importance regardless of the profits made [15]. This requirement reflects public attitudes in the United Kingdom to the use of health care data, where there is resistance to use of public data for commercial ventures unless the research could not happen without commercial involvement [16,17].

Anonymization

Data protection law only applies to personal data—that is, data that does directly or can indirectly identify an individual [18-20]. The simple deletion of name and address is usually insufficient to constitute anonymization (it has been demonstrated that the combination of 3 pieces of data could identify 87% of US residents: 5-digit zip code, birth date, and sex) [21]. The United Kingdom Information Commissioner’s Office currently treats pseudonymized data as anonymous where it is used by a third party who does not possess the requisite key code. Truly anonymized data cannot be linked back to an individual (which means that verification of data is not possible by any means). Pseudonymized data typically has identifiers removed and replaced with a unique key code (there is also 2-way cryptography; 1-way cryptography is considered anonymized). This key code can be used to trace the data back to an individual, enabling any safety concerns to be acted upon and for data to be verified. This is the approach that the United Kingdom Care.data project on the use of NHS electronic health records for data research has been taking [22]. The GDPR will require changes in practice, as it confirms in Recital 26 that pseudonymized data must be treated as personal data (in line with the previous Article 29 Working Party opinion) [18]. That position results from the increased vulnerability of data subjects who could potentially be identified compared to the protection afforded them with true anonymisation—if the key code is hacked, then all the data can be linked to an individual once more.

Consent

Consent presumed by failure to opt-out, or change preticked boxes, will no longer be permitted (unless covered by the derogations)—consent will need to be by a “clear, affirmative action” (Article 4.11). These changes would have arguably made the abandoned Care.data project [23] illegal, despite the passage of enabling legislation that exempted general practitioners from the common law duty of confidentiality when fulfilling their contractual duties to pass on health care data. Care.data relied on an opt-out for legitimacy [22]. The exercise of this opt-out was not straightforward. The numbers opting out far exceeded the estimates and the capacity for the Health and Social Care Information Centre (now NHS Digital) to process in a timely manner. The problems included omission of those who opted out from calls for NHS screening programs, even though this was not the intention of those exercising this right. NHS Digital currently relies on pseudonymization, which the GDPR states is categorized as a matter of law as personal data. It is not entirely clear whether or not third parties without access to the key code could treat pseudonymized data as anonymized (as is currently the case in the United Kingdom). Key codes are a potential vulnerability due to accidental or malicious disclosure, which is one of the justifications for pseudonymized data being classified as personal data. There are no clear indications that there are no future plans to use NHS patient data for research.

Dame Fiona Caldicott reviewed arrangements because of the widespread concerns related to consent [22], and her report led to the cancellation of the Care.data project [23]. The particular issues that were identified include the lack of information about Care.data that made exercising an opt-out an opaque process, the inadequate mechanisms for opting, and the failure of protection for rights and access to the NHS for those who opt out.

The risk of re-identification in the future is impossible to quantify precisely because it cannot be predicted what information will become public [24]. However, as with biobanks, the risks to individuals are lesser compared with studies of medical interventions [25]. Therefore authorization by research ethics committees is acceptable practice, with the requirement that opt-outs be respected unless there are exceptional circumstances.

Although the GDPR comes into force in mid-2018, researchers need to prepare now for the changes it will bring to long-term epidemiological studies. In particular, the categorization of pseudonymized data as personal will require action in some jurisdictions such as the United Kingdom and Greece [26]. The necessary accommodations will require an investment of resources, but this will hopefully ensure that subjects continue to have trust in the integrity of their health care data and the medical research community [27]. The GDPR may still apply should the United Kingdom cease to become a member state of the EU either because the United Kingdom is a member of the EEA or because the United Kingdom retains these instruments as law at least for the short term [28].

Although audit and research are treated differently in law, the boundaries between the 2 activities are blurred [29]. Audit is directly relevant to the monitoring and improvement of quality of health care; therefore, it is included as a primary use of data—Recitals 52-54 and Article 9.2 (h) and (i) of the GDPR make this clear. Audit and health care management are a primary use of health care data, and research is a secondary use—that is, it is a use different from the originally declared purpose (although it is designated a compatible purpose within the GDPR but only for nonsensitive data). If an audit compares health care systems to discover which is most effective, this can also be categorized as research as the practices are not compared to a gold standard, and there is a hypothesis being generated or even tested by finding associations. The recent furor surrounding the Royal Free Trust project in conjunction with Google DeepMind illustrates the debate over the distinction of audit from research [30-32].

Data Sharing

Dame Fiona Caldicott affirmed in her 2013 report on information governance that “The duty to share can be as important as the duty to protect patient confidentiality” [33].

Data sharing within the EU should not be obstructed because of differences in data protection law under the principles of the Digital Single Market and Article 1(2) of the Data Protection Directive. Data portability and data sharing is an issue with health care data [34], which the European Patients Smart Open Services (epSOS) project attempted to address [35]. The GDPR addresses data portability under Article 20, stating that the data subject has the right to receive their data in an appropriate format without hindrance and for data to be transferred between data controllers where technically feasible. The Bundestag is currently considering an eHealth bill with the same aim of improving portability of data [36]. This will facilitate the ability of patients to move between health care providers without unnecessary duplication of tests.

Conclusions

The Digital Single Market aims for improved data sharing across the EU, which will facilitate cross-border health care and research. Harmonization will be improved under the GDPR with a concomitant raising of standards for some countries, although there is still room for national differences according to the reasonable expectations of different publics. This advance makes cross-border projects more easily ethically justifiable and more feasible [37]. The requirements for anonymization have not been changed, except to clarify that pseudonymized data must still be considered as personal data. The GDPR will facilitate medical research, *except* where it is research not considered in the public interest. In that case, more demanding requirements for anonymization will entail either true anonymization or consent. It is likely there will be more projects that require either consent or authorization, since many projects currently use pseudonymization. There is still an unresolved issue over third parties with access to pseudonymized data.

Acknowledgments

This work has been funded by AEGLE project, Horizon 2020 ICT/2014/1 grant.

Authors' Contributions

Both authors contributed to the analysis of legal issues and the writing of the manuscript.

Conflicts of Interest

None declared.

References

1. General Data Protection Regulation 2016/679.: European Union; 2016. URL: http://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2017-02-04] [WebCite Cache ID 6o1uJtTsc]
2. LRDP Kantor, Centre for Public Reform. Comparative study on different approaches to new privacy challenges in particular in the light of technology developments. 2010. URL: http://ec.europa.eu/justice/policies/privacy/docs/studies/new_privacy_challenges/final_report_en.pdf [accessed 2017-02-04] [WebCite Cache ID 6o1upqGlT]
3. Veerus P, Lexchin J, Hemminki E. Legislative regulation and ethical governance of medical research in different European Union countries. J Med Ethics 2013 May 10;10:1. [doi: 10.1136/medethics-2012-101282]
4. Reform of EU data protection rules.: European Commission URL: http://ec.europa.eu/justice/data-protection/reform/index_en.htm [accessed 2017-02-04] [WebCite Cache ID 6o1vKzZ9R]
5. Statement by Vice-President Andrus Ansip at the press conference on the adoption of the Digital Single Market Strategy.: European Commission URL: http://europa.eu/rapid/press-release_SPEECH-15-4926_en.htm [accessed 2017-02-04] [WebCite Cache ID 6o1vbsQnB]
6. Marr B. The 5 V's of Big Data Internet.: Data Science Central; 2015. URL: http://www.datasciencecentral.com/profiles/blogs/the-5-v-s-of-big-data-by-bernard-marr [accessed 2017-02-04] [WebCite Cache ID 6o1vwOM95]
7. Analysis: Research and the General Data Protection Regulation - 2012/0011(COD) Internet.: Wellcome Trust; 2016. URL: https://wellcome.ac.uk/sites/default/files/new-data-protection-regulation-key-clauses-wellcome-jul16.pdf [accessed 2017-02-04] [WebCite Cache ID 6o1w8tR9K]
8. Stevens L. The proposed Data Protection Regulation and its potential impact on social sciences research in the UK. European Data Protection Law Review 2015;1(2):97-112.
9. Simon C, L'heureux J, Murray J, Winokur P, Weiner G, Newbury E, et al. Active choice but not too active: public perspectives on biobank consent models. Genet Med 2011 Sep;13(9):821-831 [FREE Full text] [doi: 10.1097/GIM.0b013e31821d2f88] [Medline: 21555942]
10. Hofmann B. Broadening consent—and diluting ethics? J Med Ethics 2009;35:129. [doi: 10.1136/jme.2008.024851]
11. Sheehan M. Can broad consent be informed consent? Public Health Ethics 2011;4(3):226-235. [doi: 10.1093/phe/phr020]
12. Laurie G. Governing the spaces in-between: law and legitimacy in new health technologies. In: Flear ML, Farrell A, Hervey TK, Murphy T, editors. European Law and New Health Technologies. Oxford: Oxford University Press; 2013:193.
13. Steinbekk K, Myskja B, Solberg B. Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? Eur J Hum Genet 2013;21:897-902. [doi: 10.1038/ejhg.2012.282]
14. Handyside v United Kingdom.: European Court of Human Rights; 1976. URL: http://hudoc.echr.coe.int/eng [accessed 2017-02-15] [WebCite Cache ID 6oIhGWzAg]
15. Haddow G, Laurie G, Cunningham-Burley S, Hunter K. Tackling community concerns about commercialisation and genetic research: A modest interdisciplinary proposal. Soc Sci Med 2007;64:272-282.
16. Aitken M. SHIP public engagement: summary of focus groups.: Scottish Health Informatics Programme; 2011. URL: http://www.academia.edu/2702142/SHIP_Public_Engagement_Summary_of_Focus_Groups [accessed 2017-02-04] [WebCite Cache ID 6o1y81sOm]
17. Ipsos MORI. The one-way mirror: Public attitudes to commercial access to health data. London: Wellcome Trust; 2016. URL: https://www.ipsos-mori.com/researchpublications/publications/1803/Commercial-access-to-health-data.aspx [accessed 2017-02-04] [WebCite Cache ID 6o1yFukoQ]
18. Opinion 4/2007 on the concept of personal data. Brussels: European Commission; 2007. URL: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf [accessed 2017-02-04] [WebCite Cache ID 6o1yPd6De]
19. Grubb A. Breach of confidence: anonymised information. R. v. Department of Health ex parte Source Informatics Ltd. Med Law Rev 2000;8(1):115-120. [Medline: 11787501]
20. Common Services Agency v Scottish Information Commissioner. 2008. URL: https://www.publications.parliament.uk/pa/ld200708/ldjudgmt/jd080709/comm-1.htm [accessed 2017-02-04] [WebCite Cache ID 6o1yv6syC]
21. Sweeney L. Simple demographics often identify people uniquely. Pittsburgh: Carnegie Mellon University; 2000. URL: http://dataprivacylab.org/projects/identifiability/paper1.pdf [accessed 2017-02-04] [WebCite Cache ID 6o1z6bf2W]
22. care.data hangs on engagement. 2015. URL: http://www.digitalhealth.net/2015/10/caldicott-care-data-hangs-on-engagement/ [accessed 2017-02-04] [WebCite Cache ID 6o1zGxIZu]
23. NHS England to close care.data programme following Caldicott Review.: National Health Executive; 2016. URL: http://www.nationalhealthexecutive.com/Health-Care-News/nhs-england-to-close-caredata-programme-following-caldicott-review [accessed 2017-02-04] [WebCite Cache ID 6o1zYZB6Q]
24. Anonymisation: managing data protection risk code of practice. London: Information Commissioner's Office; 2012. URL: https://ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/ [accessed 2017-02-05] [WebCite Cache ID 6o3QttPbL]
25. Laurie G, Jones K, Stevens L, Dobbs C. A review of evidence relating to harm resulting from uses of health and biomedical data.: Nuffield Council on Bioethics; 2014. URL: http://nuffieldbioethics.org/wp-content/uploads/A-Review-of-Evidence-Relating-to-Harms-Resulting-from-Uses-of-Health-and-Biomedical-Data-FINAL.pdf [accessed 2017-02-05] [WebCite Cache ID 6o3R6X2yU]

26. Data protection and research in the European Union.: European Forum for Good Clinical Practice; 2015. URL: http://www.efgcp.eu/downloads/DP%20and%20Research%20in%20EU_HD_Final_06%2010%2015.pdf [accessed 2017-02-05] [WebCite Cache ID 6o3RAFMW8]
27. Carter P, Laurie G, Dixon-Woods M. The social licence for research: why care.data ran into trouble. J Med Ethics 2015:1. [doi: 10.1136/medethics-2014-102374]
28. Mason R. Theresa May's "great repeal bill": what's going to happen and when?.: The Guardian; 2016 Oct 02. URL: https://www.theguardian.com/politics/2016/oct/02/theresa-may-great-repeal-bill-eu-british-law [accessed 2017-02-05] [WebCite Cache ID 6o3RSMQSB]
29. Wade D. Ethics, audit, and research: all shades of grey. BMJ 2005 Feb 26;330(7489):468-471 [FREE Full text] [doi: 10.1136/bmj.330.7489.468] [Medline: 15731146]
30. Hodson H. Google knows your ills. New Scientist 2016 May;230(3072):22-23. [doi: 10.1016/S0262-4079(16)30809-0]
31. Shah NR, Seger AC, Seger DL, Fiskio JM, Kuperman GJ, Blumenfeld B, et al. Improving acceptance of computerized prescribing alerts in ambulatory care. J Am Med Inform Assoc 2006;13(1):5-11 [FREE Full text] [doi: 10.1197/jamia.M1868] [Medline: 16221941]
32. ICO probes Google DeepMind patient data-sharing deal with NHS Hospital Trust Internet.: ComputerWeekly.com; 2016. URL: http://www.computerweekly.com/news/450296175/ICO-probes-Google-DeepMind-patient-data-sharing-deal-with-NHS-Hospital-Trust [accessed 2017-02-05] [WebCite Cache ID 6o3Rl0VHd]
33. Caldicott F. Information: To share or not to share? The Information Governance Review. London: Department of Health; 2013. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf [accessed 2017-02-05] [WebCite Cache ID 6o3SQDyrv]
34. Kish L, Topol E. Unpatients—why patients should own their medical data. Nat Biotechnol 2015 Sep;33(9):921-924. [doi: 10.1038/nbt.3340] [Medline: 26348958]
35. Cross-border health project epSOS: what has it achieved?.: European Commission; 2014. URL: https://ec.europa.eu/digital-single-market/en/news/cross-border-health-project-epsos-what-has-it-achieved [accessed 2017-02-05] [WebCite Cache ID 6o3SaMf6q]
36. Act on secure digital communication and applications in the health care system (E-Health Act).: Federal Ministry of Health; 2015. URL: http://www.bundesgesundheitsministerium.de/en/health/e-health-act.html [accessed 2017-02-05] [WebCite Cache ID 6o3ShEd3a]
37. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D, et al. Ethics review for international data-intensive research. Science 2016 Mar 25;351(6280):1399-1400 [FREE Full text] [doi: 10.1126/science.aad5269] [Medline: 27013718]

Abbreviations

DPD: Data Protection Directive
EEA: European Economic Area
epSOS: European Patients Smart Open Services
EU: European Union
GDPR: General Data Protection Regulation
NHS: National Health Service

Edited by G Eysenbach; submitted 07.12.16; peer-reviewed by E Afari-kumah, B Knoppers; comments to author 04.01.17; revised version received 19.01.17; accepted 21.01.17; published 24.02.17.

Please cite as:
Rumbold JMM, Pierscionek B
The Effect of the General Data Protection Regulation on Medical Research
J Med Internet Res 2017;19(2):e47
URL: http://www.jmir.org/2017/2/e47/
doi:10.2196/jmir.7108
PMID:28235748

Original Paper

Exploring the Measurement Properties of the eHealth Literacy Scale (eHEALS) Among Baby Boomers: A Multinational Test of Measurement Invariance

Lynn Sudbury-Riley[1*], PhD; Mary FitzPatrick[2*], PhD; Peter J Schulz[3*], PhD

[1]Management School, University of Liverpool, Liverpool, United Kingdom

[2]Waikato Management School, University of Waikato, Hamilton, New Zealand

[3]Institute of Communication and Health, Faculty of Communication Sciences, University of Lugano, Lugano, Switzerland

[*]all authors contributed equally

Corresponding Author:
Lynn Sudbury-Riley, PhD
Management School
University of Liverpool
Chatham Street
Liverpool, L697ZH
United Kingdom
Phone: 44 151795 ext 2553
Fax: 44 1517953001
Email: l.sudbury-riley@liverpool.ac.uk

Abstract

Background: The eHealth Literacy Scale (eHEALS) is one of only a few available measurement scales to assess eHealth literacy. Perhaps due to the relative paucity of such measures and the rising importance of eHealth literacy, the eHEALS is increasingly a choice for inclusion in a range of studies across different groups, cultures, and nations. However, despite its growing popularity, questions have been raised over its theoretical foundations, and the factorial validity and multigroup measurement properties of the scale are yet to be investigated fully.

Objective: The objective of our study was to examine the factorial validity and measurement invariance of the eHEALS among baby boomers (born between 1946 and 1964) in the United States, United Kingdom, and New Zealand who had used the Internet to search for health information in the last 6 months.

Methods: Online questionnaires collected data from a random sample of baby boomers from the 3 countries of interest. The theoretical underpinning to eHEALS comprises social cognitive theory and self-efficacy theory. Close scrutiny of eHEALS with analysis of these theories suggests a 3-factor structure to be worth investigating, which has never before been explored. Structural equation modeling tested a 3-factor structure based on the theoretical underpinning to eHEALS and investigated multinational measurement invariance of the eHEALS.

Results: We collected responses (N=996) to the questionnaires using random samples from the 3 countries. Results suggest that the eHEALS comprises a 3-factor structure with a measurement model that falls within all relevant fit indices (root mean square error of approximation, RMSEA=.041, comparative fit index, CFI=.986). Additionally, the scale demonstrates metric invariance (RMSEA=.040, CFI=.984, ΔCFI=.002) and even scalar invariance (RMSEA=.042, CFI=.978, ΔCFI=.008).

Conclusions: To our knowledge, this is the first study to demonstrate multigroup factorial equivalence of the eHEALS, and did so based on data from 3 diverse nations and random samples drawn from an increasingly important cohort. The results give increased confidence to researchers using the scale in a range of eHealth assessment applications from primary care to health promotions.

(J Med Internet Res 2017;19(2):e53) doi:10.2196/jmir.5998

KEYWORDS

health literacy; eHealth literacy; eHEALS; baby boomers; health information; measurement invariance

Introduction

The importance of health literacy for health status is well recognized. The American Medical Association, for example, found that health literacy has a stronger impact on health status than several sociodemographic variables [1] and is crucial in empowering patients to play a more active role in their own health care [2-4]. The Alliance for Health and the Future illustrates the significance of health literacy when describing it as an essential life skill for individuals, a public health imperative, an essential part of social capital, and a critical economic issue [5].

Health information is one of the most frequently sought topics on the Internet [6-8]. Consequently, in today's networked environment, electronic health resources are becoming increasingly vital in terms of overall health literacy [9,10]. New technologies that open up a myriad of eHealth applications and communications channels are revolutionizing the ways in which health information is accessed and used by both providers and patients, promising enhancement of quality of care [11] and marking a shift as patients convert from passive recipients to active consumers [7]. eHealth literacy, which is "the use of emerging information and communication technology, especially the Internet, to improve or enable health and health care" [12] (pg 267), is therefore a crucial area of study to understand and enhance the ways in which patients access and use eHealth information.

One measurement tool that is receiving increasing attention in eHealth studies is the eHealth Literacy Scale (eHEALS) [13]. A systematic review of tools to measure eHealth literacy identified 8 different measurement techniques. Noteworthy, however, is that only 1 of these techniques, eHEALS, appears in studies other than the one for which it was designed. Indeed, of 53 published articles, 45 used eHEALS [14]. Clearly, eHEALS is rapidly becoming the accepted standard way to measure eHealth literacy.

However, while there are extensive investigations pertaining to overall health literacy, the eHealth literacy construct and its psychometric properties remain understudied [15-17]. One further review of 19 health literacy instruments, including eHEALS, led to the conclusion that there are insufficient reliability assessments of data collected using health literacy scales. In fact, a key finding of this appraisal was that "limited empirical evidence exists on the reliability and construct validity of health literacy measures. This raises uncertainty about the accuracy of data being produced in relation to health literacy levels at an individual and population level" [18] (pg 367). Unsurprisingly, on this basis came a call for further research.

A further noteworthy omission from current knowledge pertaining to eHEALS is the lack of established measurement invariance. Measurement invariance, which simply means equivalence of measures, is a prerequisite before making any meaningful comparisons between different groups [19]. Indeed, too often researchers assume that an instrument developed for one culture or population automatically measures the same construct across another culture or population. However, without the establishment of measurement invariance, group comparisons are not valid or meaningful [20]. Hence, it is crucial for any scale used extensively across different nations, cultures, and groups to demonstrate measurement invariance. Developed in Canada, the scale has since been used extensively in studies with very different samples and in different cultures, including North America [21], Europe [22], and Asia [23].

Ebbinghaus [24] contended that nation-state formation, international cooperation, and easy availability of data have resulted in some countries being overrepresented (or indeed underrepresented) in many analyses. Consequently, research conducted in 1 country (usually a North American country) is assumed to be relevant to other countries, irrespective of differences in cultural and social forces. This study is part of a larger piece of research into eHealth. The choice of countries emerged from consideration of their very different rankings on health care system performance and their systems of health care provision, in the expectation that patients experiencing these different levels of services, choice, and standards would have different eHealth behaviors. What is important in this study is the basis on which we selected these countries. The Commonwealth Fund [25] ranks countries on the basis of major performance indicators on multiple health care dimensions. The analysis incorporates the views of patients and physicians pertaining to their health care systems, as well as data from the World Health Organization and the Organisation for Economic Co-operation and Development. A private foundation that promotes a good health care system to improve access, quality, and efficiency, the Commonwealth Fund supports independent research on health care issues and focuses primarily on the most vulnerable people in society (low-income people, the uninsured, minorities, young children, and elderly people). The foundation supports independent research on health care issues [25]. The ranking system used by the Commonwealth Fund [25] emerges from analysis of 80 different items pertaining to 6 main dimensions of health care: quality care (effective care, safe care, coordinated care, patient-centered care), access (cost-related problem, timeliness of care), efficiency, equity, healthy lives, and health expenditure per capita.

The United Kingdom was an obvious choice to include in the research because it ranks first overall in the league tables provided by the Commonwealth Fund [25]. While there are still major crisis points with the UK National Health Service [26,27], nevertheless the United Kingdom is ranked first across 8 of the 11 performance areas, including all of the quality-of-care indicators and the efficiency indicator. At the other extreme, ranking bottom overall, is the United States. The United States differs most notably from other industrialized nations in its lack of universal health coverage, but also ranks behind most other countries on key performance indicators pertaining to health outcomes, quality of care, and efficiency of health care delivery. Between these two extremes lies New Zealand, a country where its residents benefit from a public health system that is free or low cost due to heavy government subsidies [28], and where performance rankings are high for health measures such as effective care and coordinated care, but which lags behind many other countries in safety and equity. Notably, New Zealand is a country where eHEALS has never before been used. Hence, the inclusion of such disparate nations in this study is an important contribution to knowledge. Table 1 provides the rankings for each country in the major dimensions and subdimensions of health care provision provided by the Commonwealth Fund [25].

Table 1. Commonwealth Fund rankings of health care provision by country.

Dimension or subdimension	New Zealand	United Kingdom	United States
Quality care	4	1	5
Effective care	2	1	3
Safe care	9	1	7
Coordinated care	2	1	6
Patient-centered care	6	1	4
Access	7	1	9
Cost-related problems	6	1	11
Timeliness of care	6	3	5
Efficiency	3	1	11
Equity	10	2	1
Healthy lives	9	10	11
Health expenditure per capita (US $)	3182	3405	8508
Overall ranking	7	1	11

While the 3 countries we selected are vastly different in terms of the Commonwealth Fund health care rankings, they are nevertheless all western countries in which cultures may not differ to the extent that perhaps eastern and western nations may. Nevertheless, comparison between the 3 countries on the major cultural dimensions of national culture [29] reveal that, while they are similar in terms of high indulgence (people in high indulgence societies generally exhibit a willingness to realize their impulses and desires with regard to enjoyment of life, viewing leisure time as important, and spending money as they wish), masculinity (society is driven by competition, achievement, and success rather than caring for others and quality of life), and individualism (self-image is determined by "I" rather than "we," and personal fulfillment is important), there are some rather large differences. These differences are most notable in terms of long-term orientation, a cultural dimension that measures short-termism and quick solutions over preparing for the future. The latter dimension seems particularly important in terms of health care planning for future generations.

Baby boomers (born between 1946 and 1964) are the focus of this study. Projections suggest that this cohort will place major strains on health care systems in each of these chosen nations [30-32]. Rapid population aging and a steady increase in human longevity are leading to one of the greatest social, economic, and political transformations of all time [33]. Globally, life expectancy has increased by almost 20 years over five decades, and the profundity of this demographic change affects many economic and social areas, including health care. As longevity increases, age-related diseases such as dementia, cardiovascular disease, arthritis, osteoporosis, and type 2 diabetes will place greater demands on health care providers. Hence, in an increasingly technology-driven society, eHealth literacy is a crucially important area of study [34,35]. Many baby boomers are both technologically proficient and increasingly taking a greater role in their own health care [36]. Indeed, baby boomers have a marked difference in social attitudes in comparison with the generation that preceded them, with very different attitudes expressed in certain consumption choices, including bodily maintenance, diet, and exercise [37].

However, statistics show that baby boomers are not particularly healthy. Compared with previous generations, there is a higher prevalence of obesity, alcohol consumption, hypertension, and diabetes among baby boomers in the United States [38]; the vast majority of British boomers have at least one medical condition requiring regular medical care, with only 1 in 6 being condition-free [39]; and few doubt the significant impact that aging is predicted to have on New Zealand's health care expenditure [40]. Interestingly, the 3 countries under study rank at the bottom in terms of healthy lives (Table 1). One of the performance indicators for healthy lives is healthy life expectancy at age 60 years, and while individual ranking data for this indicator are not provided, it nevertheless gives an insight into the health-related conditions facing the baby boomers under study.

Our study therefore addressed 2 important issues. First, it answered the call for further research to examine the eHEALS, and did this through the use of structural equation modelling to examine its underlying structure. Then, by establishing full measurement invariance, our study validated eHEALS using samples of baby boomers selected from the United States, the United Kingdom, and New Zealand. We begin with a brief overview of the eHEALS and then synthesize the diverse studies that have used the scale. We then argue for the need to establish measurement invariance, before detailing the procedures used to obtain it across these diverse nations. We conclude with a discussion of the implications for future research and practice.

The eHEALS

Norman and Skinner [41] developed the lily model of eHealth literacy. The lily model depicts 6 core skills or literacies, each represented by an overlapping lily petal that feeds the pistil, which is eHealth literacy. These 6 core skills constitute 2 components. Table 2 outlines this classification of components and provides an overview of each of the core skills.

Table 2. Components of eHealth literacy lily model.

Component	Description
Analytic components: involving skills applicable to a broad range of information sources and contexts	
Traditional literacy	Ability to read text, understand written passages, and speak and write a language coherently
Information literacy	Understand how information is organized on the Internet, how to search for it, and how to use it
Media literacy	Ability to place information in a social and political context so as to understand how different media forms can shape the conveyed message
Context-specific components: situation-specific skills	
Computer literacy	Ability to use computers to solve problems
Science literacy	Ability to place health research findings in an appropriate context, thus understanding the research processes involved in knowledge creation
Health literacy	Ability to read, understand, and act on health information

Shortly after disseminating the lily model of eHealth literacy, Norman and Skinner [13] published the eHEALS, which comprises 8 items designed to “measure consumers’ combined knowledge, comfort, and perceived skills at finding, evaluating, and applying electronic health information to health problems” (pg 1). Norman and Skinner [13] reported sound scale development procedures, describing a process whereby they used the 6 core skills depicted in their lily model to compile an initial pool of items from which “an iterative process of item reduction was used to create an instrument that could be easily deployed within a variety of settings and contexts” (pg 3). This iterative process of item reduction and modification comprised reviews by faculty colleagues, a consumer group with developing literacy skills, and a large pilot test, resulting in the 8-item eHEALS shown in Table 3.

Table 3. eHealth Literacy Scale (eHEALS) scale items.

Item number	Description
1	I know what health resources are available on the Internet
2	I know where to find helpful health resources on the Internet
3	I know how to find helpful health resources on the Internet
4	I know how to use the Internet to answer my questions about health
5	I know how to use the health information I find on the Internet to help me
6	I have the skills I need to evaluate the health resources I find on the Internet
7	I can tell high-quality health resources from low-quality health resources on the Internet
8	I feel confident in using information from the Internet to make health decisions

Even from a cursory glance at the scale, it is clear that each item does not relate solely to 1 skills dimension. Rather, though it is not explicit either in the items themselves or in the published scale development article [13], it seems that embedded into each item are several core literacy skills. Item 1, “I know what health resources are available on the Internet,” is perhaps reflecting traditional and computer literacy, while item 7 could incorporate traditional, information, media, science, and health literacies. It is important to note that Norman and Skinner [13] did point out that the eHEALS does not measure the skills directly, but rather is a “measure of consumer’s perceived skills and comfort with eHealth” (pg 5).

Developed and used in further studies in Canada [42,43], the eHEALS has since been used in many countries and cultures across the globe, including the United States [34,44-54], Australia [55], Germany [56], Greece [57], Israel [58], Indonesia [59], Japan [60], the Netherlands [61,62], Norway [63], Portugal [64], Switzerland and Italy [22], Singapore [23], South Korea [65], and Taiwan [66,67], and is being used in an ongoing health intervention study in the United Kingdom [68], although results from this latter study are not yet available. The eHEALS has also been used with a wide variety of samples, including schoolchildren and adolescents [13,45,52,56,64,66,67], parents [48,69], university students [23,42], adults comprising different age groups of a wide age range [16,58,60] and adults comprising solely older generations [34,43,54], as well as veterans [46,70], patients [44,49-51,53,71], caregivers [47], and health service providers [21,59]. The scale has been used with very small (<100) sample sizes [34,42,43,45,59], as well as studies comprising several thousand respondents [48,58,60,66]. Researchers have found eHEALS to be useful for measuring perceptions of eHealth literacy to ascertain skills and training gaps [42] and to measure the success of intervention studies [34,53,68]. The scale has also been beneficial in explaining willingness to adopt personal health record technology [51]. Perhaps even more importantly, though the scale measures self-perceptions of eHealth literacy, higher scores on the scale have indicated good health behaviors, including the likelihood of undergoing cancer screening [60], as well as eating a balanced diet and taking physical exercise [72].

Clearly, eHEALS is becoming an established and well-accepted scale with which to measure eHealth literacy, used across very different studies with a wide range of research questions and a great deal of diversity in terms of sample profiles. However, often the scale is used without due consideration of its validity and reliability. It has been noted that the eHEALS construct does not appear to fully reflect the 6 different types of health literacy [18]; the representativeness of the results from smaller studies has been questioned [73]; and previous authors have noted that the validity of eHealth literacy in general [74], and the eHEALS instrument in particular [62], require further study. Moreover, the original scale authors did note that the eHealth lily model has its roots in social cognitive theory and self-efficacy theory [41]. However, despite their claim that detailed descriptions of these theories appear in their earlier publication [13], there is no explicit mention of these theories or how they were used to develop neither their eHealth literacy definition nor their eHEALS measurement instrument.

Validity and Reliability of eHEALS

Much of the burgeoning research that has used the eHEALS did so without consideration of the factorial validity of the construct. Of those studies that did examine the measurement properties of the instrument, most used principal components factor analysis [13,50,54,67]. Recently, 1 study examined the construct validity of eHEALS by first using an exploratory components analysis, which extracted 1 factor from 2 different convenience samples. Analysis then turned to further scrutiny of the scale using the Rasch model, which, in addition to providing details about the perceived difficulty of items, provides reliability statistics to estimate how well an instrument separates individuals on the construct. The study concluded that “eHEALS is a reliable and consistent measurement tool for perceived measurement of eHealth literacy. An exploratory factor analysis showed that items loaded on a single factor solution, thereby supporting the criterion of unidimensionality” [75] (pg 11).

However, while exploratory factor analysis such as principal components analysis is very useful for reducing a large number of items to a more manageable amount, a “confirmatory factor analysis of a multiple-indicator measurement model…affords a more rigorous evaluation of unidimensionality according to the constraints imposed by internal and external consistency” [76] (pg 189). Only 2 studies that we know of have used the more complex and sophisticated structural equation modelling to construct a confirmatory factor analysis (CFA) of the eHEALS. The first, conducted in Japan, entailed translation of eHEALS into Japanese [60,77] with CFA used to build a good-fitting model comprising a single factor. The second, a German study [56], compared a single-factor model to a 2-factor model. Of the 2 German alternatives, the 2-factor model was a superior fit, suggesting that the eHEALS is not unidimensional, as claimed in much previous literature, most of which has tended to use principal components analysis. However, as these authors themselves admit, the results of the 2-factor model clearly still did not indicate a well-fitting model because several important indices “indicated a poor model fit” (pg 33). Indeed, even in the better-fitting model, the root mean square error of approximation (RMSEA) was greater than 1.0, which indicates a poor-fitting model [78,79], while the comparative fit index (CFI) of .914 and the Tucker-Lewis index (TLI) of .874 are clearly not close to the .95 needed for a well-fitting model [80].

Noteworthy is that in each of the studies that used CFA, eHEALS was translated into a different language from the English in which it was originally designed. When translated, scale items can take on different meanings, and these nuances can affect perceived meanings for respondents [81,82]. The majority of health information on the Web is not only in English but developed from an English-as-a-first-language cultural perspective, and the ramifications of this appear to be far greater than for English speakers of different ethnic origins [83]. Indeed, in their original presentation of the lily model [41], Norman and Skinner commented on the fact that the overwhelming content of the Web is in English and suggested that English speakers therefore not only are more likely to find eHealth resources that are relevant to their needs, but are also more likely to find eHealth resources that they can understand. Undoubtedly, then, more research needs to examine the unidimensionality of eHEALS in an English-language context.

Importantly, to our knowledge, no previous study has examined the measurement properties of eHEALS in terms of its use with multigroups. To make comparisons between groups, measurement invariance needs to be established. Measurement invariance, or measurement equivalence, is a check to establish that a scale measures the same trait dimension, in the same way, when administered to 2 different groups [84]. Measurement invariance therefore checks that different groups (based on sex, ethnicity, nationality, or any other individual differences) respond to a measurement instrument in similar ways. Too often, researchers make assumptions about measurement equivalence, yet violations of measurement equivalence threaten fundamental interpretations of results [28]. Hence, measurement invariance is essential for testing a theory successfully in different cultural settings [19]. Without such evidence, findings “are at best ambiguous and at worst erroneous” [85] (pg 78). A standard scale, particularly one that exhibits measurement invariance, is a potentially valuable research tool for comparative and longitudinal research purposes in a variety of nations in order to create new theories or test existing hypotheses [86].

There is a growing body of international research that focuses on identifying the antecedents and impact on behavior of the eHEALS. Previous studies have examined the correlates of eHEALS in terms of antecedents such as sociodemographic characteristics [35,44], living arrangements [44], medical conditions and health status [35,44], and frequency of Internet use [44]. Additionally, some studies have attempted to measure behavioral correlates; for example, eHEALS has been described as a marker for consuming more information [87], basic Internet use [62] and using the Internet specifically for health care and lifestyle information [16,23,66], predicting postmedical visit online health information seeking [49], patient willingness to adopt a personal health record [51], and the likelihood of undergoing cancer screening [60]. A growing number of studies are also making comparisons between groups. For example, past research has made direct comparisons of eHEALS scores between different groups on the basis of various sociodemographic variables [44,64,66], and users and nonusers of Web 2.0 for health information [35]. Research has also used eHEALS to identify groups with low and high eHealth literacy and made behavioral comparisons based on these groups [22]. Establishment of measurement invariance of the scale would be a useful contribution to knowledge because measurement invariance is needed to ensure group comparisons are valid and meaningful [20]. Such groups can comprise any distinguishing measure, so to make a comparison of, say, males and females drawn from the same population, measurement invariance of a scale should be checked. This research makes that contribution.

CFA models should test a hypothesis based on a strong theoretical and empirical foundation [88]. As previously discussed, from a theoretical perspective, close scrutiny of the health literacies that make up the lily model (Table 2) and the 8 eHEALS items (Table 3) clearly shows that eHEALS does not reflect the 6 core skills depicted in the lily model. Indeed, this observation appears in previous literatures [56]. Hence, it is not easily apparent how to decide on the number of factors to test in a model based solely on the items in the lily model from which Norman and Skinner [13] claim eHEALS emerged. Norman and Skinner did, however, claim that the "foundations of the eHealth literacy concept are based in part on social cognitive theory and self-efficacy theory which promote competencies and confidence as precursors to behavior change and skill development" [13] (pg 2). It should be noted, however, that although their assertion that these theories are described in detail in their article published that same year [41], this claim does appear to be an overstatement, as there is in fact very little detail pertaining to these theories explicitly in their published work. What these authors did, however, is explain that eHEALS is based on the premise that the core skills or literacies in the lily model (Table 2) are not static and can be improved with intervention and training. In fact, they explained that literacy is as much a process as it is an outcome. It is here that social cognitive theory is apparent in their work, as social cognitive theory is based on a model of causation where behavior, environmental influences, and personal factors (which include cognitive, affective, and biological factors) all interact and influence each other [89]. Hence, rather than the lily model, here we used the underlying theories to eHEALS, namely social cognitive theory and self-efficacy theory [41], to attempt to develop a hypothesis upon which a measurement model can be tested.

The root of social cognitive theory is the concept of reciprocal determinism, where 3 factors—person, environment, and behavior—are interlinked [89]. The individual learns from experiences and the environment, which incorporates external social contexts. Responses to this learning and the environment affect the individual's behavior and therefore their ability to achieve goals. As Bandura [89] stressed, diversity in psychobiological origins, experiential conditions, and behavior results in substantial individual differences in what individuals can and cannot do. This theory therefore makes perfect sense as a foundation to eHEALS, given that individuals differ greatly in their competences pertaining to the literacies depicted in the lily model.

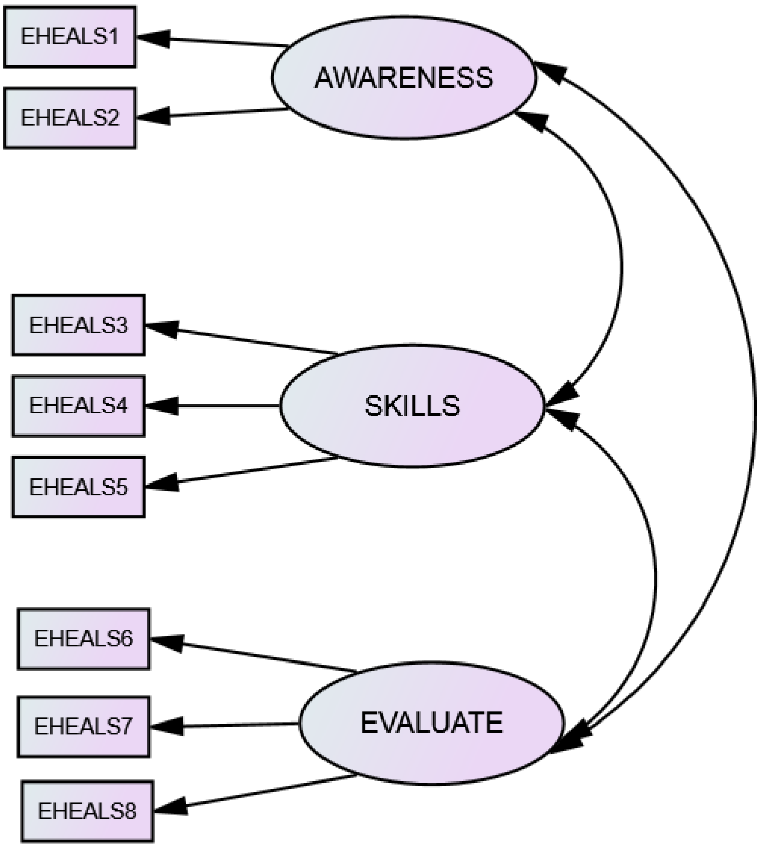
It is clear that eHEALS measures an individual's perceived skills as opposed to actual skills. An important influence in the personal dimension of the reciprocal model of social cognitive theory is self-efficacy, as this can directly influence self-motivation. Self-efficacy relates to self-belief and confidence; hence, self-efficacy is not to do with the skills a person has, but rather what that person believes they can achieve with those skills. Self-doubt and negativity can lead to failure, while self-belief and confidence can lead to an increase in effort and persistence until success is realized. Hence, self-efficacy can lead to restructuring of goals, including either lowering standards or setting higher goals to achieve even greater things, all based on the individual's perceived capabilities [90].

Attempting to apply these theories to the eHEALS is not straightforward at first glance. Nevertheless, it is relatively easy to identify those items that relate to self-efficacy. Items 6 ("I have the skills I need to evaluate…"), 7 ("I can tell high-quality….from low-quality…"), and 8 ("I feel confident in using…") all appear to pertain to a belief and confidence in one's own evaluation skills to effectively use health resources and information. However, keeping in mind that some previous empirical evidence suggests that eHEALS is neither a single-factor structure nor a 2-factor structure [56], the remaining items require close scrutiny to identify potential groupings. This close scrutiny reveals a difference between items 1 and 2, which both pertain to an awareness of what resources and information are available on the Internet, and items 3-5, which all pertain to the "how" in terms of how to find and how to use these resources. In other words, items 1 and 2 relate to an awareness of Internet health resources, items 3-5 related to the skills needed to access them, and items 6-8 relate to the self-belief that one can effectively evaluate them.

These 3 groupings do, in fact, relate to social cognitive theory in that social and technological changes affect life experiences to different degrees among different individuals [89]. Hence, knowledge of such social and technological innovations (various levels of awareness and learning about health resources on the Internet), which are reflected in items 1 and 2 of the scale, are clearly influenced by environmental factors that affect exposure to different sources of information pertaining to Internet health resources. Then, the skills needed to access these Internet health resources, which comprise items 3-5, are affected by modelling, instruction, and social persuasion in the environment. Clearly there is a behavioral element here, and such skills are a response to environmental stimuli, as well as being affected by personal factors such as internal dispositions, motivation, and biological properties that impose constraints on capabilities. This reciprocity is a key aspect of social cognitive theory [91].

Finally, self-efficacy is clearly apparent in the remaining items (items 6-8), as these items reflect an individual's self-perception of the skills needed to fully utilize the eHealth information attained on the Internet. Of course, the individual's environment and previously learned knowledge will influence the levels of self-belief that the individual holds, which is in line with the reciprocal nature of social cognitive theory. Figure 1 shows the resulting 3-factor model to be tested. Factor 1 pertains to awareness (knowledge of what resources are available and where they are), factor 2 pertains to the skills and behavior needed to access them, and factor 3 pertains to believing one has the ability to evaluate them once accessed.

Figure 1. eHealth Literacy Scale (EHEALS) 3-factor model. Factor 1 pertains to awareness of what resources are available and where they are (items 1 and 2 in the scale), factor 2 pertains to the skills and behavior needed to access them (items 3-5), and factor 3 pertains to believing one has the ability to evaluate them once accessed (items 6-8).



Methods

Instrument

The original eHEALS was developed at a time before the rise of social media [9]. Extensive social networking opportunities, as well as advances in technology such as Web 2.0, change the landscape in terms of how consumers interact with health information [92]. Hence, we tweaked the wording of the original scale items to incorporate "health information" as well as "health resources." This is because we felt that solely using the term "resources" may limit eHealth information search to official resource sites (eg, the American Cancer Society, Cancer Research UK, or Cancer Society NZ) and not incorporate the increasingly important electronic word-of-mouth that occurs on social media sites and online forums. Norman [9] advocated that the scale may need to be adapted, suggesting a social media subscale could perhaps enhance the current scale, while others have suggested that interactive applications would indeed enhance the eHEALS [93]. For these reasons, we added the words "and information" to several items. Informal feedback from friends, family, and colleagues when we asked them to name some "Internet health resources and information" reflected a wide perspective, in that people immediately cited search engines (usually Google) but also cited a wide variety of other sources, including online forums and Facebook support groups. For example, one person who was at the time undergoing tests for multiple sclerosis replied that he had not only studied the Multiple Sclerosis Society of Great Britain's webpages and viewed this as an important health resource, but also joined a Facebook group to learn more about how people coped with their diagnosis, and he viewed this as an informal information resource. Hence, rather than drastically changing the scale by adding items specific to social media, we hoped that information gained from social media would now be incorporated. Table 4 shows the adapted scale.

Approval to conduct the study was granted by the ethics committees of the University of Liverpool, the University of Waikato, and the University of Lugano.

Table 4. Adapted eHealth Literacy Scale (eHEALS).

Item number	Description
1	I know what health resources and information are available on the Internet
2	I know where to find helpful health resources and information on the Internet
3	I know how to find helpful health resources and information on the Internet
4	I know how to use the Internet to answer my questions about health
5	I know how to use the health information I find on the Internet to help me
6	I have the skills I need to evaluate the health resources and information I find on the Internet
7	I can tell high-quality health resources and information from low-quality health resources and information on the Internet
8	I feel confident in using information from the Internet to make health decisions

In addition to the tweaked eHEALS, because the study is part of a larger piece of research into eHealth, the survey contained questions pertaining to information search and usage such as sources of health information used (including interpersonal sources such as friends and family, as well as formal health information sources such as nonprofit organizations and health care providers), perceived advantages of using Internet eHealth sources (eg, 24-hour accessibility, convenience, anonymity), and perceived usefulness of Internet eHealth resources in comparison with information provided by health care providers (with a Likert-type scale ranging from “much less useful” to “much more useful”). In addition, the questionnaire contained a battery of sociodemographic variables, including age (measured via year of birth), sex (male or female), marital or relationship status (married; widowed; divorced; separated; in a domestic partnership or civil union; single, but cohabiting with a significant other; and single, never married), work status (employed full-time, employed part-time, retired, unemployed, homemaker, on government or state benefit, student or in training, other: please specify), and educational attainment (university degree; vocational training, eg, trade apprenticeship, professional qualification, college qualification; high school; less than high school).

Sample

In each country, we commissioned a commercial organization to survey randomly selected baby boomers. A prerequisite for completing the survey was that respondents (1) had to be born between 1946 and 1964 and (2) had used the Internet to search for health information in the last 6 months. Each organization was instructed to collect data from at least 250 baby boomers, and therefore the first respondents were included in the survey before the survey was closed; hence, the surveys were open for less than 2 days in each country. Prior to completing the survey, respondents were informed of its purpose (an international research project studying the use of the Internet to search for and share health information), its academic nature, how the data would be stored (password-protected secure university drives) and for how long, and the length of the survey, which typically took 20 minutes to complete. This procedure resulted in 996 usable questionnaires. There were no missing data, as a “not applicable” option was given to suitable questions, and while respondents were able to review and change their answers, they were unable to submit incomplete questionnaires.

Data Analysis

To further check the psychometric properties of the eHEALS, we conducted a series of CFAs using IBM SPSS Amos 20 (IBM Corporation). We used standard global model fit indices with well-known fit guidelines. Hence, we used the RMSEA, which is a popular measure of fit in structural equation modeling and is now recognized as one of the most informative criteria in structural equation modeling [94]. We adhered to the guidelines suggested by Hu and Bentler [80]; therefore, RMSEA values of .00 to .05 indicate a close or good fit, .05 to .08 a fair fit, .08 to .10 a mediocre fit, and over .10 a poor fit. Other fit indices that we used to assess the models were the CFI and TLI, both of which should be close to .95 [80]. The Akaike information criterion (AIC) is a fit statistic used to compare 2 models, with smaller values indicating better fit [95].

Additionally, we used probability of close fit (PCLOSE) to test the hypothesis that RMSEA is good in the population, testing the null hypothesis that RMSEA is no greater than .05 [94]. In other words, PCLOSE is an additional test of model fit, and this result indicates a close fit. Data analysis also included the use of Hoelter critical N, which is another fit statistic that differs from the others used here in that it focuses directly on the adequacy of the sample size, rather than the fit of the model. A “value in excess of 200 is indicative of a model that adequately represents the sample data” [94] (pg 83).

Steenkamp and Baumgartner [85] contended that multigroup CFA is the most powerful and versatile approach to testing for cross-national invariance and offered a sequential testing procedure for doing so. We followed this procedure here. Measurement invariance comprises 3 levels: configural, metric, and scalar. Each level is an increasingly stringent test of multigroup invariance. Consequently, we constructed a multigroup measurement model and tested it first for configural invariance, which provided a baseline model for comparisons of subsequent tests for invariance. Testing the pattern of salient (nonzero) and nonsalient (zero or near zero) loadings defined the structure of the measurement instrument [85]. In other words, the purpose of the test of configural invariance was to explore the basic structure of the construct and check that participants from different groups conceptualized the constructs in the same way [20]. Simply put, did respondents, irrespective of their cultural or national heritage, employ the same conceptual framework [96] when answering the questions that make up the eHEALS?

Configural invariance does not, however, mean that the respondents in different nations reacted to the scale items in the same way. To compare item scores meaningfully across nations, and thus have confidence in observed item differences being indicative of cross-national differences in the underlying construct, metric invariance is required. Indeed, for a scale to be useful in larger studies that examine structural relationships with other constructs cross-nationally, metric invariance is needed. [85]. Metric invariance checks that the scale is measured in the same way across groups, in that not only do different groups respond to scale items in the same way, but also the strength of the relations between items and their underlying construct is the same across groups [20].

In practice, most researchers focus on the 2 preceding and most fundamental steps, which are tests of configural and metric invariance [97]. There may be some projects, however, where researchers want to compare means and, to do this, the scale needs to exhibit scalar invariance. Scalar invariance implies that cross-national differences in the means of the observed items are due to differences in the means of the underlying constructs [85,98], and therefore indicates that the latent means can be meaningfully compared across groups [20]. Scalar invariance tests whether, in addition to the factor loadings, the intercepts are also the same, which implies that cross-national differences in the means of the observed items are due to differences in the means of the underlying constructs [98].

Results

Preliminary Analysis

Table 5 provides a profile of the sample by country.

Table 5. Sample profile by country (N=996).

Characteristics	United Kingdom (n=407)	New Zealand (n=276)	United States (n=313)	Total
Sex, n (%)				
Male	192 (47.2)	141 (51.1)	163 (52.1)	496 (49.8)
Female	215 (52.8)	135 (48.9)	150 (47.9)	500 (50.2)
Age in years, mean (SD)	59.6 (5.15)	61.3 (5.78)	60.3 (5.35)	60.3 (5.43)
Work status, n (%)				
Working full-time	132 (32.4)	82 (29.7)	84 (26.8)	298 (29.9)
Working part-time	63 (15.5)	54 (19.6)	32 (10.2)	149 (15.0)
Retired	130 (31.9)	67 (24.3)	114 (36.4)	311 (31.2)
Unemployed/welfare	35 (8.6)	42 (15.2)	27 (8.6)	104 (10.4)
Homemaker	35 (8.6)	12 (4.3)	20 (6.4)	67 (6.7)
Other	12 (2.9)	19 (6.9)	36 (11.5)	67 (6.7)
Educational attainment, n (%)				
Less than high school	3 (0.7)	0 (0.0)	8 (2.6)	12 (1.2)
High school	158 (38.8)	89 (32.2)	59 (18.8)	305 (30.6)
College/practical/technical/occupational	148 (36.4)	101 (36.6)	101 (32.3)	350 (35.1)
University degree	98 (24.1)	86 (31.2)	145 (46.3)	329 (33.0)

Table 6 provides the mean eHEALS item scores by country. While the purpose of this study was not to compare the countries in question in terms of eHealth literacy (that will be done elsewhere), noteworthy is that even a cursory glance at Table 6 reveals that US respondents had higher scores than their New Zealand and UK counterparts. We do not know whether this was due to the overall higher educational attainment of the US sample (Table 5), perceptions of poorer health care provision (Table 1), or other reasons. Across all 3 countries, the corrected item-total correlations revealed no low values (all were >.635) and it was not possible to obtain a higher alpha score by deleting any item. In all 3 nations, Cronbach alpha results were very high (.931 for the United Kingdom, .917 for the United States, and .910 for New Zealand). Indeed, with medical researchers being urged to be more critical when reporting alpha values [99], alphas this high (>.90) may suggest redundancies or that the construct being measured is too specific [100]. Hence, our analysis turned to further investigation using CFA.

Table 6. Mean eHealth Literacy Scale (eHEALS) item scores by country.

Item	United States		United Kingdom		New Zealand	
	Mean	SD	Mean	SD	Mean	SD
1	3.81	0.76	3.67	0.77	3.56	0.81
2	3.91	0.71	3.78	0.71	3.70	0.77
3	4.01	0.68	3.80	0.71	3.88	0.66
4	3.96	0.77	3.83	0.72	3.81	0.68
5	3.89	0.73	3.71	0.74	3.73	0.70
6	3.62	0.94	3.47	0.82	3.37	0.93
7	3.61	0.85	3.48	0.87	3.28	0.93
8	3.66	0.79	3.50	0.88	3.39	0.94
Cronbach alpha	.917		.931		.910	

Confirmatory Factor Analysis

The first step in testing for discriminant validity of a model structure with multiple latent factors is to reject the possibility of a single-factor structure [101]. Table 7 details these single-factor CFA results.

Table 7. eHealth Literacy Scale (eHEALS) confirmatory factor analysis by country: single-factor structure.

Country	n	χ^2	*df*	*P* value	RMSEA[a]	PCLOSE[b]	AIC[c]	CFI[d]	TLI[e]
United Kingdom	407	379.003	20	<.001	.210	<.001	411.003	.864	.809
New Zealand	276	263.140	20	<.001	.210	<.001	295.140	.833	.767
United States	313	199.218	20	<.001	.169	<.001	231.218	.896	.854

[a] RMSEA: root mean square error of approximation.

[b] PCLOSE: probability of close fit.

[c] AIC: Akaike information criterion.

[d] CFI: comparative fit index.

[e] TLI: Tucker-Lewis index.

The data did not fit the 1-dimensional model well. In addition to significant chi-square values (χ^2_{20}=379.003, *P*<.001 for the United Kingdom; χ^2_{20}=263.140, *P*<.001 for New Zealand; and χ^2_{20}=199.218, *P*<.001 for the United States), the RMSEA values of .210 for the United Kingdom and New Zealand and .169 for the United States fell outside the guidelines [78,79] proposing that values less than .05 indicate a good fit, values ranging from .05 to .08 reflect a reasonable fit, values between .08 and .10 indicate a mediocre fit, and values greater than .10 reflect a poor fit. Likewise, the CFI and TLI should be close to .95 [80], yet fell well below the cutoff point suggested for these indices in all 3 nations.

Our analysis then turned to examination of the hypothesized 3-factor model, using the UK data. Testing for factorial equivalence encompasses a series of hierarchical steps that begins with the determination of a baseline model for each group separately [94]. The first step, then, was to establish a baseline model from 1 of the samples. We chose the UK data simply because the UK sample comprises the largest number of respondents. While the 3-factor model revealed a much better fit to the 1-dimensional model, examination of the modification indices suggested improvement through the pairing of error terms associated with eHEALS items 2 and 3. One possible method effect that can trigger error covariance is a high degree of overlap in item content [94]. The high Cronbach alpha scores presented in Table 6 do of course suggest such redundancy [100]. Scrutiny of items 2 and 3 did reveal a degree of overlap, in that item 2 asks respondents if they know where to find resources, while item 3 asks them if they know how to find these resources. Clearly, to some people, there is not much difference in the meaning of these questions. Given the apparent overlap in the content of these items, and the high Cronbach alphas, which had already suggested some redundancy between scale items, we respecified the 3-factor model to include these correlated errors, and analysis moved from confirmatory to exploratory mode.

The RMSEA of .066 was within the range for a reasonable-fitting model, the CFI of .989 and the TLI of .981 far exceeded the recommended minimum values of .95, and the AIC of 84.174 shows a dramatic improvement on the previous model. Examination of the standardized residuals revealed none to exceed the threshold of 2.58 [102]; indeed, the highest standardized residual was 1.102 between eHEALS5 and eHEALS8, with all other standardized residuals falling below 1. In sum, the respecified 3-factor model fitted the UK data well.

Measurement Invariance

For the scale to be useful in multinational research, measurement equivalence is needed; without evidence of invariance, conclusions based on the scale “are at best ambiguous and at worst erroneous” [85] (pg 78). The next goal, then, was to examine the basic meaning and structure of the construct cross-nationally, to establish whether the scale is conceptualized in the same way across countries. Before moving to analysis of multinational invariance, however, Byrne [94] recommended testing the model separately in each group as the first step toward multigroup CFA. Table 8 gives the goodness-of-fit indices for each nation (including the UK data for comparative purposes). All samples demonstrated indices falling within the boundaries outlined above. Therefore, the model fit was acceptable for all countries.

Table 8. eHealth Literacy Scale (eHEALS) confirmatory factor analysis by country: 3-factor structure.

Country	n	χ^2	*df*	*P* value	RMSEA[a]	AIC[b]	CFI[c]	TLI[d]
United Kingdom	407	44.174	16	<.001	.066	84.174	.989	.981
New Zealand	276	40.651	16	.001	.075	80.651	.983	.970
United States	313	43.529	16	<.001	.075	83.529	.984	.971

[a] RMSEA: root mean square error of approximation.

[b] AIC: Akaike information criterion.

[c] CFI: comparative fit index.

[d] TLI: Tucker-Lewis index.

We then constructed a multigroup measurement model (based on the final 3-factor model) and tested it first for configural invariance. Table 9 shows the results of this and subsequent analyses. The fit indices of the configural model (χ^2_{48}=128.363, *P*<.001, RMSEA=.041, CFI=.986) indicate that the model cannot be rejected, which led to the conclusion that the specification of the items that index the 3 factors of eHEALS are configurally invariant for the 3 nations under study.

Table 9. Measurement invariance of the eHealth Literacy Scale (eHEALS) across New Zealand, the United States, and the United Kingdom.

Model	χ^2	*df*	*P* value	RMSEA[a]	PCLOSE[b]	$\Delta\chi^2$	Δ *df*	Significance	CFI[c]	ΔCFI	Critical N	
											.05	.01
1) Configural invariance	128.363	48	<.001	.041	.954	N/A[d]	N/A	N/A	.986	N/A	505	571
2) Metric invariance	149.262	58	<.001	.040	.983	20.899	10	.022	.984	.002	512	573
3) Scalar invariance	203.237	74	<.001	.042	.971	74.874	26	<.001	.978	.008	466	515

[a] RMSEA: root mean square error of approximation.

[b] PCLOSE: probability of close fit.

[c] CFI: comparative fit index.

[d] N/A: not applicable.

Table 9 also presents the results of the metric invariance analysis, when all factor loadings are constrained equally across all 3 groups. Despite the fact that metric invariance is often difficult to achieve [97], although the chi-square change between the configural and the metric model is nonsignificant, the ΔCFI of .002 is well below the proposed cutoff point of .01 [103], suggesting that the measurement model is completely invariant. This means result provides strong evidence that the eHEALS is ready to use, with a degree of confidence, in the different countries under study.

Indeed, the scale is now ready for exploring and testing structural relationships, which is the most important application for most researchers. Despite the fact that full invariance is often difficult to achieve [97], as Table 9 shows, further analyses demonstrated the eHEALS to exhibit scalar invariance; hence, analysis can include direct comparisons of mean scores. Indeed, both the “excessively stringent” [94] (pg 220) test of invariance resulting in a significant value in the change in chi-square (74.874, Δdf=26, *P*<.001), and the ΔCFI of .008 was below the .01 cutoff point [103]. Hence, despite potential social or cultural differences, the scale is unaffected. For each model, the RMSEA closeness of fit (ie, PCLOSE) far exceeds the minimum recommended *P* value of .05 [104], and Hoelter critical N at both the .05 and .01 values are greater than 200.

Despite not checking for normality prior to analysis, it is noteworthy that the data indicated no departure from normality, as evidenced by no rescaled β2 values exceeding 7 [105]. Table 10 provides these rescaled β2 values. However, there was some suggestion of multivariate kurtosis. Consequently, bootstrapping using 2000 bootstrap samples, none of which was unused, revealed only very small differences between the maximum likelihood-based estimates and the bootstrap-based estimates (Table 10). Moreover, no confidence intervals included zero (Table 11). Thus, there were no substantial discrepancies between the results of the bootstrap analysis and the original analysis, and the interpretations of the results presented above are without fear that departure from multivariate normality has biased the calculation of parameters [106].

Table 10. Rescaled β2 values and differences in maximum likelihood estimates and bootstrap estimates in the revised eHealth Literacy Scale (eHEALS) for New Zealand, the United States, and the United Kingdom.

eHEALS item	Rescaled β2 values			Differences in maximum likelihood estimates and bootstrap estimates		
	UK	NZ	US	UK	NZ	US
1	0.741	0.238	1.415			
2	2.011	1.304	1.850	.008	0.013	0.007
3	2.110	1.997	2.924			
4	1.986	1.296	2.181	0.002	0.008	0.006
5	1.709	0.746	0.934	0.003	0.008	0.015
6	0.189	0.016	0.297			
7	0.136	–0.117	–0.063	0.005	0.001	0.012
8	0.484	0.078	–0.085	0.005	0.002	0.013

Table 11. Bias-corrected bootstrap confidence intervals for the revised eHealth Literacy Scale (eHEALS).

Parameter	United Kingdom	New Zealand	United States
eHEALS1 ← awareness	0.878-1.055	0.923-1.203	0.921-1.129
eHEALS2 ← awareness	1.000-1.000	1.000-1.000	1.000-1.000
eHEALS3 ← skills	0.926-1.083	0.792-1.007	0.783-0.989
eHEALS4 ← skills	0.928-1.083	0.856-1.069	0.975-1.129
eHEALS5 ← skills	1.000-1.000	1.000-1.000	1.000-1.000
eHEALS6 ← evaluate	0.862-0.995	0.874-1.098	1.093-1.448
eHEALS7 ← evaluate	0.829-0.972	0.912-1.133	0.973-1.283
eHEALS8 ← evaluate	1.000-1.000	1.000-1.000	1.000-1.000

Finally, we tested convergent validity. First, inspection of the factor loadings presented in Table 12 revealed that all exceed the ideal of .7 [107]. Moreover, all factor loadings were positive and statistically significant.

Table 12. Standardized regression weights[a] for the revised eHealth Literacy Scale (eHEALS).

Parameter	United Kingdom	New Zealand	United States
eHEALS2 ← awareness	.919	.846	.912
eHEALS1 ← awareness	.836	.825	.857
eHEALS5 ← skills	.843	.841	.842
eHEALS4 ← skills	.877	.857	.867
eHEALS3 ← skills	.874	.832	.877
eHEALS8 ← evaluate	.843	.837	.818
eHEALS7 ← evaluate	.795	.832	.751
eHEALS6 ← evaluate	.854	.826	.730

[a] All factor loadings are positive and statistically significant.

Additionally, Table 13 presents the average variance extracted (AVE) and the construct reliability (CR) results for each country. All AVEs exceeded the cutoff of .5 [108], indicating convergent validity, and all CRs exceeded .7, indicating good reliability. Taken together, the evidence provides support for the convergent validity of the 3-construct eHEALS measurement model.

Table 13. Average variances extracted (AVE) and construct reliability (CR).

Parameter	United Kingdom		New Zealand		United States	
	AVE	CR	AVE	CR	AVE	CR
Awareness	.772	.871	.699	.822	.783	.878
Skills	.748	.898	.711	.881	.743	.897
Evaluate	.691	.870	.691	.871	.589	.811

The AVE and CR are not provided by IBM SPSS Amos software, so we calculated them using the formulae shown in Figure 2.

Figure 2. Formulae for calculating (1) average variance extracted (VE) and (2) construct reliability (CF).

$$VE = \frac{\sum_{i=1}^{n} \lambda_i^2}{n} \quad (1)$$

λ represents the standardized factor loading and *i* is the number of items.

$$CR = \frac{(\sum_{i=1}^{n} \lambda_i)^2}{(\sum_{i=1}^{n} \lambda_i)^2 + (\sum_{i=1}^{n} \delta_i)} \quad (2)$$

(δ) represents error variance terms (delta)

Discussion

Principal Findings

The finding that eHEALS comprises 3 distinct factors is a novel and important one. The eHEALS measure was developed based on the lily model of eHealth literacy, which had earlier been advanced based on social cognitive theory and self-efficacy theory [13,41]. However, to our knowledge, no previous research has ever fully examined these theories to analyze the measurement properties of the eHEALS. Despite this omission, a burgeoning body of research uses eHEALS as a measure of the extremely important concept that is eHealth literacy. As more and more of the world's population gains Internet access, and as patients increasingly expect to be active rather than passive consumers of health care services [2], the concept of eHealth will continue to grow in importance. Health care providers and researchers need a valid, reliable, and easy-to-use measurement tool with which to assess levels of perceived eHealth literacy among different groups of patients. Until now, there has been some debate about the construct validity of the eHEALS, and indeed the validity of the measurement of eHealth in general [15-18], casting doubt over subsequent results. Hopefully this study alleviates some of that doubt.

The 3 factors that emerged here are clearly based on the underlying theory on which Norman and Skinner's [41] definition of eHealth literacy is founded. From the perspective of social cognitive theory, behavioral capability refers to knowledge and skills needed to influence behavior. Additionally, human competence needs self-belief in the ability to use those skills effectively [91]. The 3-factor model presented here clearly reflects these aspects of the theory. The first factor comprises items relating to knowledge about health resources and information that are available on the Internet. The second factor relates to the skills needed to access and use the health resources and information. Finally, the third factor relates to levels of self-belief in the ability to use this information effectively.

The 3-factor structure presented here is the first to demonstrate that eHEALS does indeed relate to the social cognitive theory upon which it is founded. Future research should attempt to do the same. Indeed, all too often insufficient tests of dimensionality, reliability, and validity mar many past research studies, and it is hoped that researchers using the eHEALS measure will in future give due consideration to these crucially important dimensions of any measurement instrument. Previously, research has not given due attention to the underlying theoretical arguments for unidimensionality versus multidimensionality.

It is, of course, possible that different results have emerged here due to different populations from those that have informed past research. While eHEALS is used extensively in the United States [34,44-54], no previous study, to our knowledge, has ever used eHEALS in New Zealand, and the results of the 1 known study where it is being used in the United Kingdom [68] are not yet published. The majority of previous research that has examined the factor structure of eHEALS has used principal components analysis [13,23,50,54,64,67], rather than using CFA, which provides a much more rigorous evaluation than does principal components analysis.

Interestingly, the 2 studies that have used CFA have examined the scale after it was translated into languages other than English, the one in which it was designed. The study that found a single factor using CFA had translated eHEALS into Japanese [60,77], while the other, which found 2 factors, had translated the scale into German [56]. Hence, our study is the first to examine the factorial validity of eHEALS in the English language, in which it was originally designed. Given that translation problems can arise [81,82], it is possible that language issues have affected results in other studies. Moreover, in this study the minor tweaks to the scale in terms of insertion of the words “and information” into 5 of the items could have affected respondents’ derived meaning. Finally, the use of samples comprising solely baby boomers could have affected results. Of the 2 previous studies to use CFA, the first used a wide age range [60,77] and the second used adolescents [56]. Of course, eHEALS was originally designed using 13- to 21-year-olds [13]. We recommended that future research take these issues into account.

No previous studies that we know of have attempted multinational measurement invariance of eHEALS. Establishment of full measurement invariance is therefore another novel and important contribution. The results of a configural invariance test suggest that the respondents under study employ the same conceptual framework when answering eHEALS, despite their different cultural experiences and indeed very different experiences of health care provision.

In addition, despite it being often difficult to achieve [97], metric invariance was established. This result suggests that eHEALS is measured in the same way across these nations. Steenkamp and Baumgartner [85] (pg 82) noted that “When the purpose of the study is to relate the focal construct to other constructs in a nomological net, full or partial metric invariance has to be satisfied.” Clearly, the level of measurement invariance required for the purposes of investigating eHealth literacy in a variety of disparate nations is established.

Finally, studies may be needed to examine cross-national comparisons of the eHEALS mean scores. For such comparisons to be valid, scaler invariance needs to be established. Certainly, a cursory glance at the mean scores presented in Table 6 suggests that US baby boomers are more eHealth literate than their UK and New Zealand counterparts. While the examination and discussion of such differences is beyond the scope of this paper, it is nevertheless important to note that such comparisons can now be made legitimately, and confidence in the results of such comparisons has increased due to the establishment of full measurement equivalence.

Dolnicar and Grün [109] noted that results are only as good as the data on which they are based. Given the extensive use of eHEALS in research to date, the lack of prior research to establish measurement invariance is surprising. The results presented here therefore make an important contribution to knowledge, as “without evidence of measurement invariance, the conclusions of a study must be weak” [110] (pg 119). In establishing full measurement invariance, comprising configural, metric, and scalar invariance, this study has demonstrated that eHEALS is now ready to use with confidence in these diverse nations. Moreover, the AVE and the CR results for each nation all suggest convergent validity and good reliability. Overall, these results provide solid support for the convergent validity of the 3-factor eHEALS model.

In addition to the implications of this study for theory, these findings are important from a practical perspective. Results demonstrate that, consistent with the theory on which it was developed, eHEALS assesses self-perceptions of 3 important and distinct (though interrelated) elements of eHealth literacy: awareness of Internet health resources (items 1 and 2), the skills needed to access them (items 3-5), and the self-belief that one can effectively evaluate them (items 6-8). Hence, eHEALS can now be used to segment health consumers into distinct groups based on their scores on the scale, with corresponding intervention and training provision designed around meeting the needs of these segments. Those individuals with relatively low scores on the awareness factor would need to be offered basic training designed to address the rudimentary elements of eHealth in terms of describing and demonstrating the range of appropriate resources available and how they can be found. For people whose scores are relatively low on this factor, such training should perhaps be stand-alone and could be the foundational level of training. Once they master these skills, individuals could be offered the second level of training, designed for those people whose scores are relatively low on the skills factor. This skills training should be designed to perhaps build on basic knowledge and concentrate on developing the individual’s search and evaluation skills pertaining to eHealth resources. Finally, a third training program could be developed that concentrates on developing and building self-efficacy, to give people the self-belief that they are truly empowered patients who are able to play an active role in their own health care. Most training and educational programs incorporate levels of progression in their design, and eHealth intervention and training programs should be no different.

Practical intervention and training around eHealth is important for several reasons. eHealth has the potential to assist self-management in people with chronic health conditions, and evidence suggests that even in developed countries, half of the population with chronic health conditions have elementary navigational needs and would benefit from basic training in this area [55]. Training programs are crucial because patients with higher levels of health literacy have significantly lower anxiety levels than people with inadequate health literacy, and have fewer and shorter consultations with health care providers [111]; hence, there are economic benefits to such training programs. Improvements in ability and self-belief to access and use Internet health resources have cumulative benefits in terms of ability and willingness to use other eHealth resources such as electronic health records, patient portals, and self-management tools [74]. Thus, understanding different skill levels and needs is important for policy makers and health care providers, who could all use such information to develop correct and targeted interventions for different segments of the population. Indeed, it has even been suggested that eHealth is so important that it should be incorporated into school curricula [57]. When they first designed eHEALS in 2006, Norman and Skinner [13] claimed that the scale has the potential to identify those who may or may not benefit from referrals to an eHealth intervention or resource. Our research builds on this claim and suggests that eHEALS can be used to ascertain the type of intervention or resource that could benefit these different segments.

Limitations

The study is not without its limitations. First, while baby boomers are a justifiably important sample for health care and eHealth research, the 3-factor structure that emerged here needs to be investigated using younger samples to ensure that boomers are not unique and the 3-factor structure is indeed applicable to all age groups. Second, while the 3 nations we chose do vary a great deal in terms of health care provision rankings and to a lesser extent on some important cultural dimensions, they are nevertheless all English-speaking western countries. It has been noted that when eHEALS was translated, different factorial structures emerged. We recommend that the 3-factor model be tested in very diverse cultures (eg, eastern countries) and among non-English-speaking nations. Third, we acknowledge that the original version of eHEALS was designed before the rise in social media and Web 2.0 technology. While we made some attempt to incorporate the interactive nature of today's online environment by tweaking the scale (specifically, adding "and information" to items), the suggestion that the marginally updated version used here is sufficient to incorporate interactive resources is based solely on anecdotal evidence gained by asking family, friends, and colleagues. We recommend that a more formal study investigate the way respondents perceive the eHEALS in its revised form, as it may need to be more extensively altered, or indeed a new scale may need to be designed, in order to fully capture the myriad of interactive eHealth resources that health care consumers are now able to access.

Conclusions

The usefulness of a short, easy-to-administer scale that measures a person's perception of their eHealth literacy is beyond doubt. Indeed, the extensive use of the eHEALS across a variety of studies in countries across the globe is testimony to the urgent requirement for such an instrument. The research presented here details a more rigorous investigation of the measurement properties of the eHEALS than has previously been conducted, using CFA rather than principal components analysis. Based on social cognitive theory and self-efficacy theory, a 3-factor model was tested and confirmed.

Research often needs to make comparisons across groups or across time and, to be able to do this, a scale must demonstrate measurement invariance. Only by establishing measurement invariance can there be assurance that comparisons are valid [112]. In other words, establishing measurement invariance provides evidence that score differences across countries are a true representation of differences in the construct under study, rather than differences brought about by social and cultural factors or other such confounding variables [96]. This research has demonstrated full measurement invariance of the eHEALS among baby boomers in 3 diverse nations, meaning the scale is now ready to use with far more confidence by researchers in these nations. This research has therefore added weight to Norman and Skinner's [13] contention that the scale is a useful addition to a range of eHealth assessments, from primary care to health promotions. The identification of 3 distinct factors not only confirms the theoretical antecedents on which eHEALS was built, but also suggests that the scale can now be used to better segment health care consumers and identify different skills gaps, enabling policy makers and health care providers to design and offer tailored interventions and training programs to address such gaps.

Over 80% of baby boomers in all 3 countries under study use the Internet regularly [113-115]. Nevertheless, this cohort did not grow up using the Internet, and there may be some for whom knowledge, skills, and self-confidence around eHealth resources still lag behind the levels that perhaps exist among younger cohorts. Yet the baby boomer cohort is crucially important from an eHealth perspective because forecasts predict that this cohort is increasingly going to put major pressures on health care systems [30-32]. Importantly, Bandura [91] explains that personal factors can be altered dramatically to improve the functioning of individuals. Competency can be developed through training and guidance, which in turn can increase self-belief in capability levels. While eHealth training lessons are already available across all 3 countries we studied, the findings suggest these training programs should be built around knowledge of what health information and resources are available on the Internet, and then developing the skills needed to access them. Motivational enhancements should also be incorporated into such training to ensure an enhancement in self-belief.

In sum, this study fills an important gap in that it provides future researchers and practitioners with more faith in the eHEALS than existed previously. The scale can now be used with a degree of confidence in a variety of nations and in studies with a variety of research objectives, including the modeling of complex relationships among variables. The choices of nations and the demographic of the samples therein are also strengths of the study: all too often, scale evaluation and development comprises young (often student and often US) samples. Studies often use scales developed in a different country or culture without checking that the measure is equivalent. This study has demonstrated that eHEALS can be used with confidence across a variety of nations and cultures. This study therefore lends support for the contention that eHEALS is a valid scale with which to measure self-perceptions of eHealth literacy, a concept that is set to become even more important in the future.

Conflicts of Interest

None declared.

References

1. Flynn D, Gregory P, Makki H, Gabbay M. Expectations and experiences of eHealth in primary care: a qualitative practice-based investigation. Int J Med Inform 2009 Sep;78(9):588-604. [doi: 10.1016/j.ijmedinf.2009.03.008] [Medline: 19482542]
2. Coulter A, Ellins J. Effectiveness of strategies for informing, educating, and involving patients. BMJ 2007 Jul 7;335(7609):24-27 [FREE Full text] [doi: 10.1136/bmj.39246.581169.80] [Medline: 17615222]
3. 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]
4. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
5. Kickbusch I, Wait S, Maag D. Navigating health: the role of health literacy. London, UK: Alliance for Health and the Future; 2005. URL: http://www.ilcuk.org.uk/images/uploads/publication-pdfs/pdf_pdf_3.pdf [accessed 2016-05-05] [WebCite Cache ID 6hH1ORa1N]
6. Fox S. 80% of Internet users look for health information online. Washington, DC: Pew Internet and American Life Project; 2011 Feb 1. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2011/PIP_Health_Topics.pdf [accessed 2017-02-13] [WebCite Cache ID 6oFgOjJZ4]
7. McMullan M. Patients using the Internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
8. Bidmon S, Terlutter R. Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. J Med Internet Res 2015;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]
9. Norman C. eHealth literacy 2.0: problems and opportunities with an evolving concept. J Med Internet Res 2011 Dec;13(4):e125 [FREE Full text] [doi: 10.2196/jmir.2035] [Medline: 22193243]
10. Seçkin G, Yeatts D, Hughes S, Hudson C, Bell V. Being an informed consumer of health information and assessment of electronic health literacy in a national sample of internet users: validity and reliability of the e-HLS instrument. J Med Internet Res 2016 Jul 11;18(7):e161 [FREE Full text] [doi: 10.2196/jmir.5496] [Medline: 27400726]
11. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
12. Eng TR. eHealth research and evaluation: challenges and opportunities. J Health Commun 2002 Jul;7(4):267-272. [doi: 10.1080/10810730290001747] [Medline: 12356287]
13. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
14. Karnoe A, Kayser L. How is eHealth literacy measured and what do the measurements tell us? A systematic review. Knowledge Manage E-Learning 2015;7(4):576-600.
15. Collins SA, Currie LM, Bakken S, Vawdrey DK, Stone PW. Health literacy screening instruments for eHealth applications: a systematic review. J Biomed Inform 2012 Jun;45(3):598-607 [FREE Full text] [doi: 10.1016/j.jbi.2012.04.001] [Medline: 22521719]
16. Hu X, Bell RA, Kravitz RL, Orrange S. The prepared patient: information seeking of online support group members before their medical appointments. J Health Commun 2012;17(8):960-978. [doi: 10.1080/10810730.2011.650828] [Medline: 22574697]
17. van Deursen A, van Dijk J. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
18. Jordan JE, Osborne RH, Buchbinder R. Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. J Clin Epidemiol 2011 Apr;64(4):366-379. [doi: 10.1016/j.jclinepi.2010.04.005] [Medline: 20638235]
19. Mathur A, Barak B, Zhang Y, Lee KS. A cross-cultural procedure to assess reliability and measurement invariance. J Appl Meas 2001;2(3):241-255. [Medline: 12011509]
20. Milfont T, Fischer R. Testing measurement invariance across groups: applications in cross cultural research. Int J Psychol Res 2010;3(1):111-121.
21. Blackstock OJ, Haughton LJ, Garner RY, Horvath KJ, Norwood C, Cunningham CO. General and health-related Internet use among an urban, community-based sample of HIV-positive women: implications for intervention development. AIDS Care 2015 Apr;27(4):536-544. [doi: 10.1080/09540121.2014.980215] [Medline: 25411825]
22. Diviani N, van den Putte B, Meppelink CS, van Weert JCM. Exploring the role of health literacy in the evaluation of online health information: insights from a mixed-methods study. Patient Educ Couns 2016 Jun;99(6):1017-1025. [doi: 10.1016/j.pec.2016.01.007] [Medline: 26817407]
23. Suri V, Majid S, Chang Y, Foo S. Assessing the influence of health literacy on health information behaviors: a multi-domain skills-based approach. 2015 Presented at: Annual Health Literacy Research Conference; November 2-4, 2015; Bethesda, MD, USA.

24. Ebbinghaus B. When less is more: selection problems in large-N and small-n cross-national comparisons. Int Sociol 2005 Jun 01;20(2):133-152. [doi: 10.1177/0268580905052366]
25. Davis K, Stremikis K, Squires D, Schoen C. Mirror, mirror on the wall: how the performance of the U.S. health care system compares internationally: 2014 update. New York, NY: The Commonwealth Fund; 2014 Jun. URL: http://www.commonwealthfund.org/~/media/files/publications/fund-report/2014/jun/1755_davis_mirror_mirror_2014.pdf [accessed 2017-02-13] [WebCite Cache ID 6oFirFNAB]
26. Chakrabortty A. The Guardian. 2016 Feb 08. This NHS crisis is not economic: it's political URL: http://www.theguardian.com/society/2016/feb/08/this-nhs-crisis-is-not-economic-its-political [accessed 2017-02-13] [WebCite Cache ID 6oFj3Bcjx]
27. Donnelly L. The Telegraph. 2015 Oct 09. NHS faces biggest financial crisis 'in a generation' URL: http://www.telegraph.co.uk/news/nhs/11921381/NHS-faces-biggest-financial-crisis-in-a-generation.html [accessed 2017-02-13] [WebCite Cache ID 6oFjB9ykI]
28. New Zealand Now. Wellington, NZ: Ministry of Business, Innovation & Employment; 2017. Healthcare: New Zealand's comprehensive health system is built on Kiwis' inbuilt need to see that everyone gets 'a fair go' in life URL: https://www.newzealandnow.govt.nz/living-in-nz/healthcare [accessed 2017-02-13]
29. Hofstede G. Country Comparison. Helsinki, Finland: itim International URL: https://geert-hofstede.com/countries.html [accessed 2017-02-13] [WebCite Cache ID 6oFo5yzk1]
30. Carter T. What is costing the NHS?. London, UK: Centre for Policy Studies; 2015 Dec 07. URL: http://www.cps.org.uk/blog/q/date/2015/12/07/what-is-costing-the-nhs/ [accessed 2017-02-13] [WebCite Cache ID 6hcMSzrfx]
31. Gillies A. New Zealand Herald. 2012 Nov 15. Baby boomers outgrowing other age groups URL: http://www.nzherald.co.nz/lifestyle/news/article.cfm?c_id=6&objectid=10847534 [accessed 2017-02-13] [WebCite Cache ID 6hcMlQ1Fi]
32. Johnson D. Time Magazine. 2015 May 11. These charts show the baby boomers' coming health crisis URL: http://time.com/3852306/baby-boomer-health-charts/ [accessed 2016-05-19] [WebCite Cache ID 6hcMyPQnI]
33. United Nations Department of Economic and Social Affairs, Population Division. World population prospects: key findings & advance tables: 2015 revision. New York, NY: UN; 2015. URL: http://esa.un.org/unpd/wpp/Publications/Files/Key_Findings_WPP_2015.pdf [accessed 2017-02-13] [WebCite Cache ID 6hcO5hSEV]
34. Xie B. Effects of an eHealth literacy intervention for older adults. J Med Internet Res 2011 Nov;13(4):e90 [FREE Full text] [doi: 10.2196/jmir.1880] [Medline: 22052161]
35. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
36. KPMG International. Accelerating innovation: the power of the crowd. Global lessons in eHealth implementation. 2012 Mar. URL: https://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/accelerating-innovation/Documents/ehealth-implementation.pdf
37. Biggs S, Phillipson C, Leach R, Money A. The mature imagination and consumption strategies: age & generation in the development of a United Kingdom baby boomer identity. Int J Ageing Later Life 2008 Apr 03;2(2):31-59. [doi: 10.3384/ijal.1652-8670.072231]
38. King DE, Matheson E, Chirina S, Shankar A, Broman-Fulks J. The status of baby boomers' health in the United States: the healthiest generation? JAMA Intern Med 2013 Mar 11;173(5):385-386. [doi: 10.1001/jamainternmed.2013.2006] [Medline: 23381505]
39. Medical Research Council. Most baby boomers 'under the doctor' at retirement. London, UK: MRC; 2012 Sep 19. URL: http://www.mrc.ac.uk/news/browse/most-baby-boomers-e28098under-the-doctore28099-at-retirement/ [accessed 2016-05-19] [WebCite Cache ID 6hcPJwGHn]
40. Davey J. Health Literacy: A Hermeneutic Study of New Zealand Baby Boomers [dissertation]. Hamilton, New Zealand: University of Waikato; 2015.
41. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
42. Brown CA, Dickson R. Healthcare students' e-literacy skills. J Allied Health 2010;39(3):179-184. [Medline: 21174023]
43. Manafò E, Wong S. AAssessing the eHealth literacy skills of older adults: a preliminary study. J Consumer Health Internet 2012 Oct;16(4):369-381. [doi: 10.1080/15398285.2012.701163]
44. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
45. Ghaddar SF, Valerio MA, Garcia CM, Hansen L. Adolescent health literacy: the importance of credible sources for online health information. J Sch Health 2012 Jan;82(1):28-36. [doi: 10.1111/j.1746-1561.2011.00664.x] [Medline: 22142172]
46. Haun JN, Patel NR, Lind JD, Antinori N. Large-scale survey findings inform patients’ experiences in using secure messaging to engage in patient-provider communication and self-care management: a quantitative assessment. J Med Internet Res 2015;17(12):e282 [FREE Full text] [doi: 10.2196/jmir.5152] [Medline: 26690761]

47. Horvath KJ, Courtenay-Quirk C, Harwood E, Fisher H, Kachur R, McFarlane M, et al. Using the Internet to provide care for persons living with HIV. AIDS Patient Care STDS 2009 Dec;23(12):1033-1041. [doi: 10.1089/apc.2009.0163] [Medline: 20025513]
48. Knapp C, Madden V, Marcu M, Wang H, Curtis C, Sloyer P, et al. Information seeking behaviors of parents whose children have life-threatening illnesses. Pediatr Blood Cancer 2011 May;56(5):805-811. [doi: 10.1002/pbc.22674] [Medline: 21370415]
49. Li N, Orrange S, Kravitz RL, Bell RA. Reasons for and predictors of patients' online health information seeking following a medical appointment. Fam Pract 2014 Oct;31(5):550-556. [doi: 10.1093/fampra/cmu034] [Medline: 24963151]
50. Nahm E, Resnick B, Orwig D, Magaziner J, Bellantoni M, Sterling R, et al. A theory-based online hip fracture resource center for caregivers: effects on dyads. Nurs Res 2012;61(6):413-422 [FREE Full text] [doi: 10.1097/NNR.0b013e318270712a] [Medline: 23052420]
51. Noblin AM, Wan TTH, 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]
52. Paek H, Hove T. Social cognitive factors and perceived social influences that improve adolescent eHealth literacy. Health Commun 2012;27(8):727-737. [doi: 10.1080/10410236.2011.616627] [Medline: 22452551]
53. Robinson C, Graham J. Perceived Internet health literacy of HIV-positive people through the provision of a computer and Internet health education intervention. Health Info Libr J 2010 Dec;27(4):295-303 [FREE Full text] [doi: 10.1111/j.1471-1842.2010.00898.x] [Medline: 21050372]
54. Sheng X, Simpson PM. Seniors, health information, and the Internet: motivation, ability, and Internet knowledge. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):740-746. [doi: 10.1089/cyber.2012.0642] [Medline: 23679569]
55. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer use of “Dr Google”: a survey on health information-seeking behaviors and navigational needs. J Med Internet Res 2015;17(12):e288 [FREE Full text] [doi: 10.2196/jmir.4345] [Medline: 26715363]
56. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement. J Media Psychol 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
57. Xesfingi S, Vozikis A. eHealth literacy: in the quest of the contributing factors. Interact J Med Res 2016 May 25;5(2):e16 [FREE Full text] [doi: 10.2196/ijmr.4749] [Medline: 27226146]
58. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
59. Rachmani E, Dai HJ, Hsu CY. Internet use behavior of leprosy control program staff at Indonesia’s public health centre. Indian J Med Inform 2014;2(8):24-27.
60. Mitsutake S, Shibata A, Ishii K, Oka K. Association of eHealth literacy with colorectal cancer knowledge and screening practice among internet users in Japan. J Med Internet Res 2012;14(6):e153 [FREE Full text] [doi: 10.2196/jmir.1927] [Medline: 23149453]
61. Ossebaard HC, Seydel ER, van Gemert-Pijnen L. Online usability and patients with long-term conditions: a mixed-methods approach. Int J Med Inform 2012 Jun;81(6):374-387. [doi: 10.1016/j.ijmedinf.2011.12.010] [Medline: 22261086]
62. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
63. Antypas K, Wangberg SC. E-Rehabilitation: an Internet and mobile phone based tailored intervention to enhance self-management of cardiovascular disease: study protocol for a randomized controlled trial. BMC Cardiovasc Disord 2012;12:50. [doi: 10.1186/1471-2261-12-50] [Medline: 22776554]
64. Tomás C, Queirós P, Ferreira T. [Analysis of the psychometric properties of the Portuguese version of an eHealth literacy assessment tool]. Rev Enfermagem Referência 2014 Jun 30;4(2):19-28. [doi: 10.12707/RIV14004]
65. Cho J, Park D, Lee HE. Cognitive factors of using health apps: systematic analysis of relationships among health consciousness, health information orientation, eHealth literacy, and health app use efficacy. J Med Internet Res 2014;16(5):e125 [FREE Full text] [doi: 10.2196/jmir.3283] [Medline: 24824062]
66. Chang F, Chiu C, Chen P, Miao N, Lee C, Chiang J, et al. Relationship between parental and adolescent eHealth literacy and online health information seeking in Taiwan. Cyberpsychol Behav Soc Netw 2015 Oct;18(10):618-624. [doi: 10.1089/cyber.2015.0110] [Medline: 26375050]
67. Koo M, Norman C, Hsiao-Mai C. Psychometric evaluation of a Chinese version of the ehealth literacy scale (eHEALS) in school age children. Int Electronic J Health Educ 2012;15:29-36.
68. Thomas CL, Man M, O'Cathain A, Hollinghurst S, Large S, Edwards L, et al. Effectiveness and cost-effectiveness of a telehealth intervention to support the management of long-term conditions: study protocol for two linked randomized controlled trials. Trials 2014;15:36 [FREE Full text] [doi: 10.1186/1745-6215-15-36] [Medline: 24460845]
69. Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. J Med Internet Res 2011;13(3):e75 [FREE Full text] [doi: 10.2196/jmir.1697] [Medline: 21960017]

70. Haun JN, Lind JD, Shimada SL, Martin TL, Gosline RM, Antinori N, et al. Evaluating user experiences of the secure messaging tool on the Veterans Affairs' patient portal system. J Med Internet Res 2014;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
71. Marrie RA, Salter A, Tyry T, Fox RJ, Cutter GR. Health literacy association with health behaviors and health care utilization in multiple sclerosis: a cross-sectional study. Interact J Med Res 2014;3(1):e3 [FREE Full text] [doi: 10.2196/ijmr.2993] [Medline: 24513479]
72. Mitsutake S, Shibata A, Ishii K, Oka K. Associations of eHealth literacy with health behavior among adult internet users. J Med Internet Res 2016 Jul;18(7):e192 [FREE Full text] [doi: 10.2196/jmir.5413] [Medline: 27432783]
73. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
74. Watkins I, Xie B. eHealth literacy interventions for older adults: a systematic review of the literature. J Med Internet Res 2014;16(11):e225 [FREE Full text] [doi: 10.2196/jmir.3318] [Medline: 25386719]
75. Nguyen J, Moorhouse M, Curbow B, Christie J, Walsh-Childers K, Islam S. Construct validity of the eHealth Literacy Scale (eHEALS) among two adult populations: a Rasch analysis. JMIR Public Health Surveill 2016 May;2(1):e24 [FREE Full text] [doi: 10.2196/publichealth.4967] [Medline: 27244771]
76. Gerbing DW, Anderson JC. An updated paradigm for scale development incorporating unidimensionality and its assessment. J Marketing Res 1988 May;25(2):186-192. [doi: 10.2307/3172650]
77. Mitsutake S, Shibata A, Ishii K, Okazaki K, Oka K. [Developing Japanese version of the eHealth Literacy Scale (eHEALS)]. Nihon Koshu Eisei Zasshi 2011 May;58(5):361-371. [Medline: 21905612]
78. Browne M, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. Testing Structural Equation Models. Newbury Park, CA: Sage; 1993:136-162.
79. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1(2):130-149. [doi: 10.1037/1082-989X.1.2.130]
80. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling 1999;6(1):1-55. [doi: 10.1080/10705519909540l]
81. Harkness J, van der Vijver FJR, Mohler P. Cross-Cultural Survey Methods. Hoboken, NJ: Wiley; 2002.
82. Rode N. Translation of measurement instruments and their reliability: an example of job-related affective well-being scale. Metodoloski Zvezki 2005;2(1):15-26.
83. Werts N, Hutton-Rogers L. Barriers to achieving e-health literacy. Am J Health Sci 2013;4(3):115-120.
84. Widaman K, Reise S. Exploring the measurement invariance of psychological instruments: applications in the substance use domain. In: Bryant KJ, Windle M, West SG, editors. The Science of Prevention: Methodological Advances From Alcohol and Substance Abuse Research. Washington, DC: American Psychological Association; 1997:281-324.
85. Steenkamp J, Baumgartner H. Assessing measurement invariance in cross-national consumer research. J Consumer Res 1998 Jun;25(1):78-107. [doi: 10.1086/209528]
86. Ackroyd S, Hughes J. Data Collection in Context. New York, NY: Longman; 1981:0582480159.
87. Neter E, Brainin E. eHealth literacy: a marker for “digital divide” in health information. Rev Health Care 2012 Jul 12;3(3):145. [doi: 10.7175/rhc.v3i3.281]
88. Stevens J. Applied Multivariate Statistics for the Social Sciences. 3rd edition. Mahwah, NJ: Lawrence Erlbaum Associates; 1996.
89. Bandura A. Social cognitive theory. In: Vasta R, editor. Annals of Child Development. Volume 6. Greenwich, CT: JAI Press; 1989:1-60.
90. Reynolds A. Age-Related Differences in Self-Efficacy and the Use of e-Health Supports for Exercise Behavior in Adults [dissertation]. Gainesville, FL: University of Florida; 2013.
91. Bandura A. Organizational applications of social cognitive theory. Austn J Manage 1988 Dec 01;13(2):275-302. [doi: 10.1177/031289628801300210]
92. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. J Med Internet Res 2011 Nov;13(4):e94 [FREE Full text] [doi: 10.2196/jmir.1750] [Medline: 22094891]
93. van der Vaart R, Drossaert CHC, de Heus M, Taal E, van de Laar MAFJ. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. J Med Internet Res 2013 Feb;15(2):e27 [FREE Full text] [doi: 10.2196/jmir.2428] [Medline: 23399720]
94. Byrne B. Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming. 2nd edition. New York, NY: Routledge, Taylor & Francis; 2010:978.
95. Hu L, Bentler P. Evaluating model fit. In: Hoyle RH, editor. Structural Equation Modeling: Concepts, Issues, and Applications. Thousand Oaks, CA: Sage; 1995:76-99.
96. Wu A, Li Z, Zumbo B. Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: a demonstration with TIMSS data. Pract Assess Res Eval 2007;12(3):1-26.
97. Chen Y, Tang TL. Attitude toward and propensity to engage in unethical behavior: measurement invariance across major among university students. J Bus Ethics 2006 Jul 5;69(1):77-93. [doi: 10.1007/s10551-006-9069-6]

98. Hu C, Cheung GW. Eight general questions on measurement equivalence/invariance that bother me most. Acad Manage Proc 2008 Aug 01;2008(1):1-6. [doi: 10.5465/AMBPP.2008.33641699]
99. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd]
100. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. J Pers 1986 Mar;54(1):106-148. [doi: 10.1111/j.1467-6494.1986.tb00391.x]
101. Kline R. Principles and Practice of Structural Equation Modelling. 3rd edition. New York, NY: Guildford; 2011:978-971.
102. Jöreskog K, Sorbom D. New developments in LISREL. 1987 Mar Presented at: The National Symposium on Methodological Issues in Causal Modeling; March 1987; Tuscaloosa, AL, USA.
103. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling 2002 Apr;9(2):233-255. [doi: 10.1207/S15328007SEM0902_5]
104. Jöreskog K, Sörbom D. LISREL 8 User's Reference Guide. Lincolnwood, IL: Scientific Software International; 1996.
105. West S, Finch J, Curran P. Structural equation models with non-normal variables: problems and remedies. In: Hoyle RH, editor. Structural Equation Modeling: Concepts, Issues, and Applications. Thousand Oaks, CA: Sage; 1995:56-75.
106. Garson G. Structural Equation Modeling. Ashboro, NC: Statistical Associates Publishers; 2015.
107. Hair J, Black W, Babin B, Anderson R. Multivariate Data Analysis: A Global Perspective. 7th edition. Upper Saddle River, NJ: Prentice Hall; 2009.
108. Ping RA. On assuring valid measures for theoretical models using survey data. J Business Res 2004 Feb;57(2):125-141. [doi: 10.1016/S0148-2963(01)00297-1]
109. Dolnicar S, Grün B. Cross‐cultural differences in survey response patterns. Int Marketing Rev 2007 Apr 24;24(2):127-143. [doi: 10.1108/02651330710741785]
110. Horn J. Comments on 'issues in factorial invariance'. In: Collins LM, Horn JL, editors. Best Methods for the Analysis of Change. Washington, DC: American Psychological Association; 1991:114-125.
111. Rowlands GP, Mehay A, Hampshire S, Phillips R, Williams P, Mann A, et al. Characteristics of people with low health literacy on coronary heart disease GP registers in South London: a cross-sectional study. BMJ Open 2013 Jan 03;3(1):3e001503 [FREE Full text] [doi: 10.1136/bmjopen-2012-001503] [Medline: 23293243]
112. van de Schoot R, Lugtig P, Hox J. A checklist for testing measurement invariance. Eur J Dev Psycho 2012 Jul;9(4):486-492. [doi: 10.1080/17405629.2012.686740]
113. Office for National Statistics. Internet Users in the UK 2016. Newport, UK: ONS; 2016. URL: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2016#main-points [accessed 2017-02-13] [WebCite Cache ID 6oFnWYHti]
114. Statistica. New York, NY: Statistica. Share of Internet users in the United States as of May 2015, by age group URL: https://www.statista.com/statistics/266587/percentage-of-internet-users-by-age-groups-in-the-us/ [accessed 2017-02-13] [WebCite Cache ID 6oFnnQlfM]
115. Techday NetGuide. Auckland, New Zealand: Techday Ltd; 2011 Nov 01. Baby boomers catching up to younger generations' technology skills URL: https://netguide.co.nz/story/baby-boomers-catching-up-to-younger-generations-technology-skills/

Abbreviations

AIC: Akaike information criterion
AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: comparative fit index
CR: construct reliability
eHEALS: eHealth Literacy Scale
PCLOSE: probability of close fit
RMSEA: root mean square error of approximation
TLI: Tucker-Lewis index

Edited by G Eysenbach; submitted 23.05.16; peer-reviewed by M Stellefson, R van der Vaart; comments to author 31.07.16; revised version received 23.10.16; accepted 30.01.17; published 27.02.17.

Please cite as:
Sudbury-Riley L, FitzPatrick M, Schulz PJ
Exploring the Measurement Properties of the eHealth Literacy Scale (eHEALS) Among Baby Boomers: A Multinational Test of Measurement Invariance
J Med Internet Res 2017;19(2):e53
URL: http://www.jmir.org/2017/2/e53/
doi:10.2196/jmir.5998
PMID:28242590

Original Paper

Readiness for Delivering Digital Health at Scale: Lessons From a Longitudinal Qualitative Evaluation of a National Digital Health Innovation Program in the United Kingdom

Marilyn R Lennon[1], BSc, PGDip, PhD; Matt-Mouley Bouamrane[1], PhD; Alison M Devlin[2], BSc (Hons), PhD; Siobhan O'Connor[3], RN, BSc, CIMA CBA; Catherine O'Donnell[2], BSc (Hons), MPH, FHEA, PhD; Ula Chetty[2], BSc, MRCGP, MBChB, MRCP; Ruth Agbakoba[2], MSc, B Sc (Hons); Annemieke Bikker[2], MSc; Eleanor Grieve[4], BA (Hons), MPH; Tracy Finch[5], BA (Hons), PhD; Nicholas Watson[6], BSc (Hons), MSc, PhD; Sally Wyke[6], FRSE, FRCGP, PhD, B Sc (Hons); Frances S Mair[2], FRCGP, DRCOG, MD

[1]Digital Health and Wellness Group, Computer and Information Sciences, University of Strathclyde, Glasgow, United Kingdom

[2]General Practice and Primary Care, Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

[3]School of Health and Social Care, Edinburgh Napier University, Edinburgh, United Kingdom

[4]Health Economics and Health Technology Assessment, Institute of Health and Wellbeing, University of Glasgow, Glasgow G12 8RZ, UK, United Kingdom

[5]Institute of Health and Society, Newcastle University, Newcastle Upon Tyne NE2 4AX, United Kingdom

[6]School of Social and Political Sciences, Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:
Frances S Mair, FRCGP, DRCOG, MD
General Practice and Primary Care
Institute of Health and Wellbeing
University of Glasgow
Horselethill Road
University of Glasgow
Glasgow, G12 9LX, UK
United Kingdom
Phone: 44 01413308317
Fax: 44 01413308317
Email: Frances.Mair@glasgow.ac.uk

Abstract

Background: Digital health has the potential to support care delivery for chronic illness. Despite positive evidence from localized implementations, new technologies have proven slow to become accepted, integrated, and routinized at scale.

Objective: The aim of our study was to examine barriers and facilitators to implementation of digital health at scale through the evaluation of a £37m national digital health program: Delivering Assisted Living Lifestyles at Scale" (dallas) from 2012-2015.

Methods: The study was a longitudinal qualitative, multi-stakeholder, implementation study. The methods included interviews (n=125) with key implementers, focus groups with consumers and patients (n=7), project meetings (n=12), field work or observation in the communities (n=16), health professional survey responses (n=48), and cross program documentary evidence on implementation (n=215). We used a sociological theory called normalization process theory (NPT) and a longitudinal (3 years) qualitative framework analysis approach. This work did not study a single intervention or population. Instead, we evaluated the processes (of designing and delivering digital health), and our outcomes were the identified barriers and facilitators to delivering and mainstreaming services and products within the mixed sector digital health ecosystem.

Results: We identified three main levels of issues influencing readiness for digital health: macro (market, infrastructure, policy), meso (organizational), and micro (professional or public). Factors hindering implementation included: lack of information technology (IT) infrastructure, uncertainty around information governance, lack of incentives to prioritize interoperability, lack of precedence on accountability within the commercial sector, and a market perceived as difficult to navigate. Factors enabling implementation were: clinical endorsement, champions who promoted digital health, and public and professional willingness.

Conclusions: Although there is receptiveness to digital health, barriers to mainstreaming remain. Our findings suggest greater investment in national and local infrastructure, implementation of guidelines for the safe and transparent use and assessment of digital health, incentivization of interoperability, and investment in upskilling of professionals and the public would help support the normalization of digital health. These findings will enable researchers, health care practitioners, and policy makers to understand the current landscape and the actions required in order to prepare the market and accelerate uptake, and use of digital health and wellness services in context and at scale.

(J Med Internet Res 2017;19(2):e42) doi:10.2196/jmir.6900

KEYWORDS

telemedicine; health plan implementation; community health services; health services research; electronic health records; instrumentation; qualitative research; diffusion of innovation; medical informatics

Introduction

It is often the case with eHealth and digital health studies that 1 single service or product is studied in a controlled setting (often a randomized control trial) to determine its effectiveness in order to proceed to roll it out at scale and make it part of routine care delivery pathways. Previous research, however, has shown that uptake and adoption are slow for digital health overall and that there may be many sociotechnical, organizational, or cultural barriers that are slowing the mainstreaming of digital health [1,2]. Over the last two decades, there has been an exponential growth in the literature describing barriers and facilitators to innovation [3] and digital health implementation [4]. This literature initially focused on examining implementation issues on a relatively small scale [5,6] but this was followed by more extensive studies looking at large scale deployments of single digital health services in particular contexts [7-11]. Issues that have been identified as barriers to implementation range from liability concerns, interoperability issues, costs, usability, misaligned incentives, and policy problems through to acceptability to patients and professionals [4,8].

The concept of eHealth readiness” has previously been highlighted [12,13] and there is increasing interest in using tools to better measure specific aspects of digital health readiness [14-16]. However, the existing evidence mainly relates to organizational readiness within the health service [17,18] or the readiness of health professionals [19-21], patients or carers for a specific single digital health service [22]. The technical, political, and market preparedness and readiness for widespread delivery of consumer-oriented digital health services which encompass and cross health-social-technological boundaries has not yet been fully explored. Consequently, a whole systems” analysis of readiness for digital health is warranted.

Delivering Assisted Living Lifestyles at Scale (dallas) was an ambitious national program conducted from May 2012 to May 2015 in the United Kingdom. The program received £37m ($55m, €50m) in funding from Innovate UK, the National Institute for Health Research (NIHR), The Scottish Government, Scottish Enterprise, and Highlands and Islands Enterprise. The dallas program aimed to develop and implement a wide range of digital health and wellness products and services to enable preventive care, self-care, and independent living at scale. One of the program’s primary goals was to stimulate the consumer market for person-centered digital technologies. It was explicitly set up as a large scale research and development program rather than a randomized clinical trial or a series of individual pilots. This was considered crucial by the program funders to begin to understand what the existing barriers to uptake and adoption of digital health at scale are and to unlock new markets and pathways to make digital health at scale a reality.

Four large multi-agency consortia (referred to as communities) called i-Focus,” Living It Up,” More Independent,” and Year Zero” were funded (see Multimedia Appendix 1 for details of each consortium). The communities were funded specifically to design, deploy, and promote awareness and uptake of a range of innovative digital health and wellness services across the United Kingdom. Innovation and ability to scale up and sustain digital health and open new routes to market were considered key to the program’s success. Unlike many other previous digital health trials or studies, the services developed were aimed at a broad sociodemographic, including children, parents, older adults, as well as the broader consumer health and fitness population living in urban or remote and rural regions. Some of the services were digital and aimed at increasing awareness, redesigning services or care pathways, as well as increasing the uptake and routinization of digital health as a whole. Details of these consortia have been reported previously [23], but Multimedia Appendix 1 briefly summarizes the wide range of stakeholders involved in each consortium and the range of digital health products, services, and activities developed and delivered via the dallas program.

Due to the variety of products and services (apps, personal health records, telecare, telehealth, wearable activity trackers, and so on) and the variety of populations and contexts, communities were measuring a range of traditional primary outcomes such as levels of engagement, perceived usability and acceptability of the products, and reduction in resource usage (such as hospitalizations). In addition to this, however, the communities were exploring ways to capture changes (positive increases) in the amount and quality of choice, contribution, community, collaboration, and connectedness that the new services created.

In this paper, we present synthesized qualitative findings from a longitudinal study of digital health design, delivery, and roll out. We examine implementation issues from different angles and with different stakeholders with a focus on what this data tells us about the readiness of different elements of the ecosystem in the United Kingdom to deliver digital health at scale. Given the current self-care agenda, the drive toward more personalized medicine [24,25], and person-centered digital health solutions [26,27], such work is timely and has the potential to make an important contribution to understanding the implementation of digital health innovations internationally. The aim of this study was to capture barriers and facilitators to implementation of digital health across a wide range of stakeholders and across time, thus allowing us to answer the question of how ready” different people, processes, and systems are for mainstreaming digital health and to identify what measures might be taken to reduce some of the existing and persistent barriers in this area. Here we present our findings and conclude with a set of 10 recommendations to address some of the key readiness barriers identified.

Methods

Aim and Design

Longitudinal qualitative and survey data were collected over 39 months (June 2012-October 2015) to help us identify and understand key barriers and facilitators experienced during the implementation journey. Tables 1 and 2 illustrate the full breadth and volume of data collected.

Table 1. Qualitative implementation dataset (interviews).

Interviews		Number of items	Number of participants	Number of pages
eHealth Implementation Toolkit (e-HIT) interviews		47	53	1134
	Interviews: Baseline e-HIT Dates: October 2012-January 2013 Participants: 7 health, 6 industry, 3 voluntary, 2 academia Dallas community: iF, LiU, Mi, YZ	17	18	247
	Interviews: Midpoint e-HIT Dates: October 2013-December 2014 Participants: 13 health, 7 industry, 3 voluntary, 1 academia Dallas community: iF, LiU, Mi, YZ	20	24	630
	Interviews: Endpoint e-HIT Dates: May-October 2015 Participants: 5 health, 5 industry, 1 voluntary sector Dallas community: iF, LiU, Mi, YZ	10	11	257
Champions		22	23	233
	Interviews: Lay champions Dates: December 2013-December 2014 Participants: 11 volunteer champions, 2 voluntary sector, 1 health service, 2 administrator, 1 other Dallas community: Mi, LiU	17	17	186
	Interviews: Digital champions Dates: March 2015 Participants: 5 voluntary champions, 1 government Dallas community: Mi	5	6	47
LiU stakeholders		32	20	545
	Interviews: LiU longitudinal interviews Dates: January 2014-January 2015 Participants: 5 health service managers; 1 health service lead Dallas community: LiU	18	6	315
	Interviews: LiU cross-sectional interviews Dates: July 2014-April 2015 Participants: 2 health prof, 6 industry, 2 voluntary sector, 2 government, 1 academia, 1 consumer user Dallas community: LiU	14	14	230
Project management or cross project		24	26	248
	Interviews: Evaluation alignment Dates: May-November 2014 Participants: 1 health service, 4 industry Dallas community: iF, Mi, YZ	5	5	11

Interviews		Number of items	Number of participants	Number of pages
	Interviews: dallas leads Dates: June 2015 Participants: 3 health service, 2 industry Dallas community: iF, LiU, Mi, YZ	5	5	46
	Interviews: Digital Health And Care Alliance Dates: March-April 2015 Participants: 1 health, 4 industry, 1 academia, 1 voluntary, 3 government Dallas community: iF	10	10	168
	Interviews: House of Memories Dates: July-October 2015 Participants: 2 patients, 2 carers, 1 industry, 1 government Dallas community: Mi	4	6	23
Subtotal		125	122	2,160

Table 2. Qualitative implementation dataset (focus groups).

Focus groups	Number of items	Number of Participants	Number of pages
Focus group : Lay champions Dates: December 2013 Participants: 8 volunteer champions Dallas community: Mi	1	8	23
Focus group: House of Memories Dates: March 2015 Participants: 4 patients, 4 carers, 1 health service and 1 government agency staff Dallas community: Mi	1	10	40
Focus group: eRedBook Dates: April 2015 Participants: 12 health service users, 9 health professional users, 2 health service; 2 industry Dallas community: YZ	2	25	71
Focus group: No delays Dates: April 2015 Participants: 4 health service users, 4 health service staff & 1 administrator Dallas community: YZ	2	9	49
Focus group: Get active Dates: May 2015 Participants: 5 users & 2 voluntary sector staff Dallas community: LiU	1	7	19
Subtotal	7	59	202

Detailed evidence was gathered from numerous stakeholders rolling out different services to enable a rich understanding of digital health readiness (see Tables 1 and 2). From this, detailed reports describing the diverse experiences of each group within their context, the process of rolling out products or services, and factors that shaped each consortium’s implementation journey were written. Cross-case analysis of communities was conducted, drawing out not only commonalities related to readiness” but also differences or alternative explanations of factors affecting readiness for digital health at the individual, organizational, or wider environmental and political level.

Sampling and Setting

Specific roles and partner organizations were identified by the research team as critical to capturing perspectives of each stakeholder group within the consortia. This was wide ranging and included health care professionals; health and social service managers; staff and volunteers from third sector organizations;

private companies designing, developing, and promoting hardware and software platforms; and academics and employees from government agencies working on guidelines and policies. Geographic locations spanned England and remote rural and urban regions of Scotland. Care was taken to include representatives from all types of organizations involved (private, public, voluntary) and to include stakeholders involved at the strategic level, the project management level, and the service design and delivery level. Progress of the program was followed longitudinally over 3 years. Interviews were undertaken with different stakeholders over this period. A subset of key individuals was subjected to repeat interviews, for example, as part of our eHealth Implementation Toolkit (e-HIT) interviews and our case study work, allowing us to access the perspectives of key implementers at different points during the program.

Data Collection and Management

Perspectives of key implementers from each community were sought via semistructured interviews. The structure of the interview was based on expanding on issues raised when using the e-HIT tool at baseline, midpoint, and end of the program [28]. This tool was developed in previous research by members of the team [28] and is designed to help promote understanding of digital health implementation issues. Focus groups were held with end users including patients, carers, consumers, health professionals, and local champions (see Table 2). Interviews and focus groups were audio-recorded with participant consent, transcribed verbatim, and anonymized. This dataset was supplemented with reflective notes and ethnographic data noted during primary data collection. Electronic and paper-based survey data collection was also undertaken to gather opinions from health professionals.

Theoretical Approach

Increasingly, in implementation research, the use of theory has been advocated in order to allow us to develop an improved understanding and explanation of why service innovations or digital health technologies become an integrated part of routine service delivery or not [29]. A range of theories have been utilized such as actor network theory [30] through to theories of organizational readiness [31].

Our evaluation was underpinned by a sociological theory, normalization process theory (NPT) [32,33], which has been used extensively to enhance understanding of how individuals and groups of people understand, integrate, and sustain new technologies, service innovations, or ways of working into everyday practice [32]. NPT has 4 core constructs (Multimedia Appendix 2) and is the underpinning for the 2 instruments used as part of the evaluation toolkit. The e-HIT [28] was used for 47 qualitative stakeholder interviews undertaken over the 3 years of the project (baseline, midpoint, and end point, see Table 1) and the NoMAD [34] questionnaire, derived from NPT constructs, was used with 48 health professional respondents involved with the No Delays Service. One community (Living it Up) was also studied more widely across stakeholders and across time as a case study.

NPT provided a consistent and coherent theoretical lens to analyze and interpret data across the program which enabled us to systematically identify themes and to provide structure to any explanation we could identify in the data for the emerging key themes. Our use of NPT as the theoretical underpinning of our analysis across the dallas program has allowed us to use NPT as a lens through which to conceptualize data at different levels, thereby taking into account wider contextual and environmental factors as well as workability issues at the individual level. It has therefore helped us to make propositions or recommendations for future large scale digital health implementation programs.

Data Analysis

Qualitative data analysis was informed by a framework approach [35], using a coding frame informed by NPT [23]. The data (interviews, project documentation, field notes) were presented as transcripts and notes and the text coded (tagged within qualitative data software called NVivo (QSR International) to generate an annotated coding book. Coding clinics were undertaken to ensure consistency of coding and shared understanding of coding constructs. A total of 8 researchers with multidisciplinary experience in health services research, clinical research, informatics, and social science were involved at coding clinics with 1 or 2 senior academics (FSM, MML) involved in all coding sessions. Care was taken to expand, collapse, and rename codes so that all were confident that the coding book was a fair and accurate reflection of the data. Integrative analyses and key mapping of emerging themes undertaken during the final phase of the project identified readiness” concepts as a key theme visible across the NPT framework. A matrix of overarching readiness themes was then coded in parallel with our NPT framework (see Table 3). Issues falling outside the NPT framework, like organizational culture,” were noted.

Ethics and Governance

This evaluation was commissioned by Innovate UK. University ethical approvals were granted for the collection of all qualitative data reported (College of Medical, Veterinary and Life Sciences, Approval: 200130141; and College of Science and Engineering, Approval: CSE01210 and CSE01096 at the University of Glasgow as well as University of Newcastle ethics approval Reference number 00555/2012). Our work was overseen by an external advisory group. Informed consent to participation was gained from all participants. Anonymity was protected by use of generic descriptors throughout.

Results

Analysis and Coding Scheme

The first significant result to report is the breadth and volume of data collected across the stakeholders via the different mixed methods. Table 1 in the methods section provides an overview that helps to contextualize where our results, discussed in the following section, emerge from. We next present an overview of the final coding structure with discussions and exemplar quotes from each emerging theme. A more detailed set of representative quotes per theme can be found in Multimedia Appendix 3.

Our analysis revealed that readiness issues were present at 3 levels (Figure 1): Macro (market, policy or governmental, wider context), meso (commercial, organizational, and infrastructure), and micro (professional, public, carers).

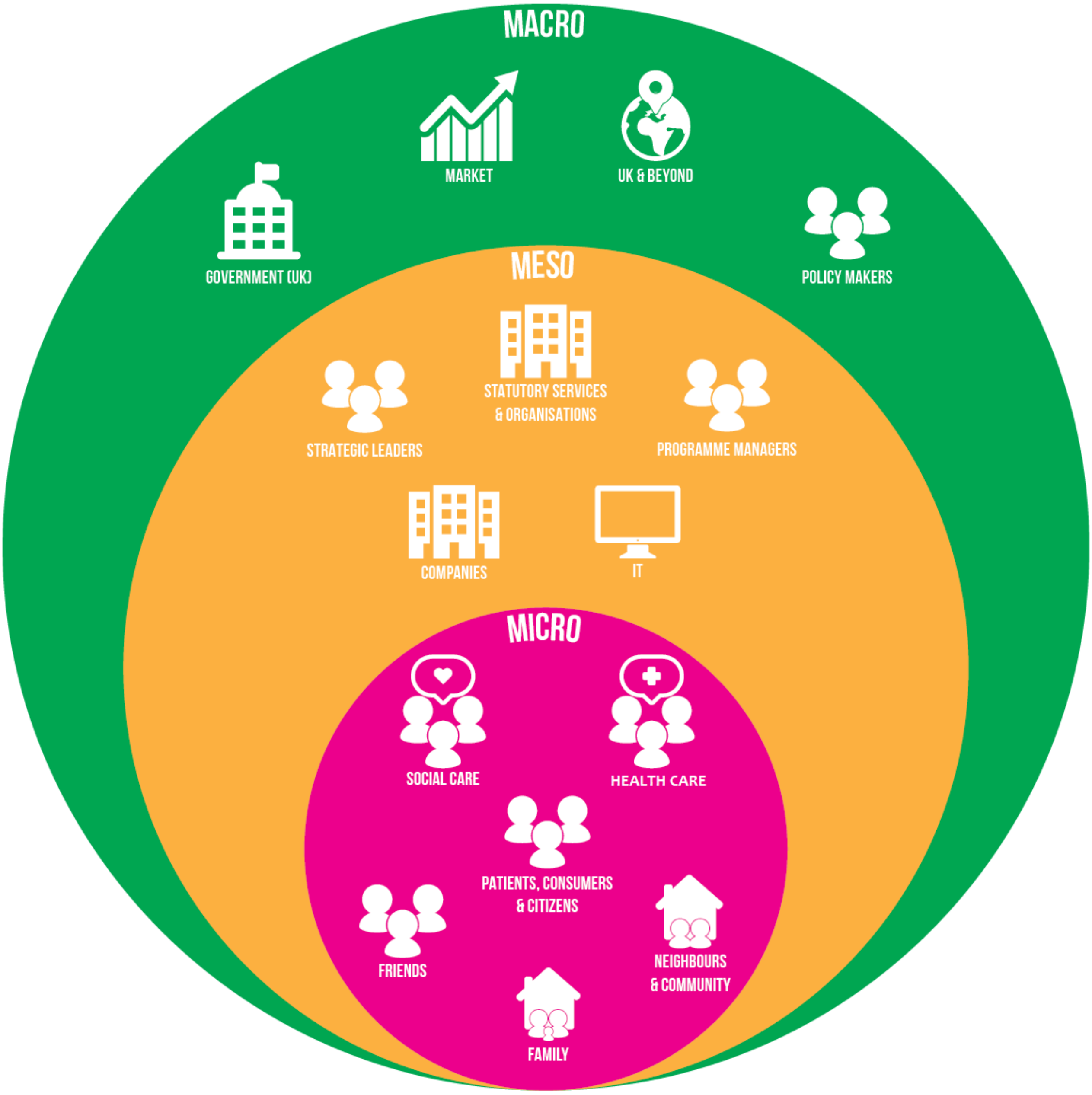
The full coding scheme is illustrated in Table 3 together with how it maps onto our NPT theoretical framework.

Table 3. Overview of readiness coding scheme mapped to normalization process theory (NPT) constructs.

Level	Theme
Macro	
Market	Interoperability (collective action) Risk or liability (coherence and collective action) Clinical endorsement (collective action)
Policy or infrastructure	National policy (collective action) Infrastructure (collective action)
Meso	
Industry	Incoherent market (coherence, cognitive participation)
National Health System (eg, National Health Service)	Information technology infrastructure (collective action) Discontinuity and organizational culture (some collective action or cognitive participation but some outside normalization process theory) Resources (collective action)
Micro	
Health professional readiness	Workload and professional confidence (collective action) Training & alignment with professional roles or identity (collective action) Access to digital resources (collective action)
Public readiness of digital health services and systems	Digital literacy and access (cognitive participation or collective action) Agency and lifestyle (coherence) Security and trust (collective action)

The key issues (those that were most prevalent in the data) under each of these main readiness” themes are now described in detail with illustrative quotations provided to support each key point (see Multimedia Appendix 3 for additional supporting quotes). The discussion section goes on to present recommendations based on these findings which we argue will accelerate the uptake, mainstreaming, and ultimate success of digital health at scale.

Figure 1. Key themes influencing readiness for digital health.



Macro-Level Readiness

As described in our overarching approach, it was considered critical to examine not just single products or geographies in localized contexts but to also use a wider lens to capture barriers or facilitators to digital health at the full ecosystem level. This includes the systems and structures socially, technically, and politically that are needed, and how ready these systems are to support and even promote digital health. Two key overarching themes were identified at the macro level: market readiness” and national policy or infrastructure readiness.” They are presented here with each of their subthemes with representative quotes from the data.

Market Readiness

One of the program’s primary goals was to stimulate the consumer market for person-centered digital technologies. This was identified as an ongoing tension and required continual navigation and negotiation during the life of the program. Four key areas contributed to this challenge at the time of writing and are described in turn here.

Interoperability

Interoperability is a key issue for digital health products and services, particularly when they have to interface and exchange data with National Health Service (NHS) and other clinical and social care information systems. To address this, 1 consortium, i-Focus, invested significant time and resource developing a reference architecture which could serve as a common technical framework to work across sectors. This consortium also established the Digital Health and Care Alliance (DHACA), a not-for-profit, member-driven organization, building understanding and sharing knowledge and expertise between small- and medium-sized enterprises and the statutory sector on how to develop and implement interoperable digital health and wellness services.

Our data revealed that interoperability was perceived as more than a technical challenge. Commercial companies often perceive open standards and interoperability as a threat to their business model, since it was not a priority in comparison with efforts to increase their own individual market share as the market develops and matures. Lack of interoperability was a clear barrier to progress.

> *...there is an element of maybe we could work more together, but where do I spend my time? Do I spend my time in a meeting discussing data integration with another company when I've got only tens of thousands users worldwide, or do I try and get into users worldwide then worry about it?* [DHACA interviewee no. 2, 2015]

Risk and Liability

Providing digital health services for people with major frailty or multimorbidity was still perceived as risky in terms of product liability issues by some commercial companies. Concerns were raised by industry partners—especially those that traditionally operate outside health and social care—about the responsibilities entailed by operating within the digital health statutory sphere.

> *It’s a young person’s market, or it’s a worried well” market. So the people who buy fitbit and pedometers. And that's great, (...) because they are responsible for the outcome of that monitoring. If you’re monitoring a heart condition, you want someone at the end of that. You don’t just want a fitbit, you want a Triage Nurse. So...I’m not entirely convinced that there is a consumer market in these technologies, actually...I think there's a hybrid market, maybe,...And I think we might be a way off, you know, finding it...it is really complex.* [DHACA interviewee No. 9, 2015]

Providing self-monitoring devices as part of health promotion was deemed acceptable but providing clinical data back to those with known health problems was perceived as a different proposition, involving significant risk. Companies often wanted to sell a technology kit,” whereas the statutory sector wanted a different type of contract,” linking the use of technology to data services and outcomes. This was a risk which companies were reluctant to take on in the current market until it was clear what the ethical guidelines and responsibilities were around collection and use of lifestyle data and was another barrier to implementation.

Clinical Endorsement

Accreditation and clinical endorsement were seen as crucial issues affecting deployment plans for digital health. Clinical endorsement could involve a single health professional endorsing it to people in their practice for example, or more likely a body of clinicians publically backing up or signing up to say that they think that the product is useful and clinically beneficial. This can still have a huge effect on digital health product success—even if that product is meant to appeal to the consumer staying well, as opposed to a patient being treated for a condition.

A purely consumer version of the eRedBook, a digital child health record created as part of the Year Zero consortium, did not prove viable initially. This was because endorsement by the relevant medical association was seen as a prerequisite to ensure uptake by lay users and health professionals; obtaining such accreditation was labor and time intensive and posed a barrier for implementation but once accreditation or endorsement was achieved, this was perceived as a potential facilitator of uptake.

> *What our experience brought us to realize is that people will only use a personal health journal around serious or long term conditions if it’s something they can engage with their clinicians on. You’re not going to persuade people to go out and buy it as a consumer product if it’s not something their clinicians will engage with them on and look at and share the information that they've been collecting.* [C3 Implementer interview, June 2015]

Accreditation and official endorsement of digital health products and services were seen as key ongoing issues likely to influence deployments and future development of digital health, and further research and policy work is required to clarify what apps and services require accreditation (and which do not) and also what such accreditation should look like.

Complexity of the Market

The digital health market proved difficult to access and navigate especially for international companies or start-ups unfamiliar with the landscape. The organization and delivery of health care is currently devolved to the 4 countries of the United Kingdom. Each individual health service is composed of a large number of heterogeneous and autonomous organizations functioning in substantially distinct ways. The interface between health and social care varies and many products and services are now more lifestyle or wellness based and not clearly and solely within the remit of either health or social care uniquely. All of this makes the UK digital health market challenging to navigate, with a lack of clear access or entry points for the retail sectors. The eclectic nature of the dallas consortia helped provide opportunities for people to connect and learn how to navigate such a complex landscape and to experiment with different models and pathways to implement and mainstream digital health that might not have been considered traditionally.

> *...the amount of red tape from the National Health Service and the...finance committees, procurement committee, it would have been very difficult to know who would be the right person to speak to so it’s kind of opened opportunities for people to get round the table and have real discussions about how they can make a difference and that’s been a really positive part of the program.* [Final e-HIT interview C2, interview 1, social care manager]

Political Readiness and National Policy

Information governance policies and legislation issues within the health and social care sectors were a recurring theme. Regulations around information governance generally are strongly embedded and well established in the UK health service due to its culture of high security, with patient confidentiality viewed as a priority. Although policy and legislation relating to data sharing has been reviewed and clear recommendations made [36], this has not yet translated to local contexts and was reported as a key barrier to deployment of the dallas program across its lifetime.

> *There is a real problem...health data is in a vault that’s owned by the National Health Service. You can’t, at the moment, view it and when you can view it, it will be a view which is not in a form that can be used by technology outside of the NHS in any real useful way. I think the biggest issue is information governance and letting people take ownership of their own data and their own risk appetite, and until that happens all we are doing is allowing the market to develop outside of the true record.* [DHACA interview 3, 2015]

The notion of sharing sensitive health data across multiple public and private organizations that do not hold the same information governance rules is fraught with difficulty. This led to a common view that information governance regulations were not currently fit for purpose.” If we are trying to change existing care paradigms, responsibilities, and data ownership for digital health, it is clear that further work is needed around specific information governance for health and wellness products that are not covered by existing clinical or statutory policy or governance. For example, accessing Twitter” or similar sites using hospital computer systems was often not permissible, which meant that integrating social networking platforms within the health care arena was problematic.

Some recent national policies were seen as positive drivers for change among certain consumer groups. For example, the experience from digital enablement activities suggested that recent social benefits reforms compelled people to sign on the Web.” This encouraged many individuals to improve their computer literacy by joining Lay” and Digital champion” programs and to engage with digital hubs so they could access their social welfare benefits. As a result, disadvantaged groups potentially at risk of digital exclusion were provided with digital skills and educated about health technologies. Policy and funding streams need to advocate and support digital inclusion and awareness-raising if digitally supported self-management is to become a reality for people on the ground accessing the services.

> *...the benefits reform has been a great carrot or stick to push people...you’ll hear stories from Digital Champions, people coming in on a Friday afternoon because they are going to get sanctioned if they don't do this online form...* [Digital champion interview, March 2015]

Infrastructure

Our data showed significant variation in national infrastructure across the United Kingdom. Those in remote and rural areas voiced concerns about inadequate Internet connectivity as a limiting factor for accessing digital health services. Health centers in urban areas also reported to lack the connectivity necessary to enable access to new digital health services being rolled out. Organizations at the local and national level clearly need to invest in information technology (IT) infrastructures if digital services are to be rolled out and supported robustly across the United Kingdom.

> *Area (X) did phenomenally well given the poor connectivity in the region you know poor WiFi and even when we 3G-enabled their tablets poor 3G signals you know, it was a hard slog of going around x centers and signing people up.* [C3 final e-HIT interview 1]

Meso-Level Readiness

At the intermediate or meso” level there were 2 main themes identified: Industry” and local health service organizational” readiness. These are related to the specific markets and organizations required to access and roll out digital health.

Industry Readiness

Digital health is constantly promoted as a potentially lucrative market. However, enticing commercial entities—who normally sell products directly to consumers—to invest in opportunities in emerging digital health, wellbeing, or social care sectors did not prove as straightforward as originally anticipated in the dallas program. This may be due in part to lack of a coherent market at the time of the program.

Lack of Market Coherence

Market stability and maturity were key themes for industry readiness. Private industries that normally operate outside of health and social care were reticent to engage with the relatively immature, digital health sector, proving less ready” to invest than anticipated. One community tried hard to engage with well-respected retailers but found it impossible to translate initial interest into actual delivery of digital health offerings. In addition, some private industries did not fully grasp what consumers wanted or required of them in terms of digital health products.

> *…I think user experience is key so start with the language of the consumer and language of the value proposition, people who are selling to consumers who are parting with their money rather than looking at the language of local authorities and health sector which is all about cost avoidance…* [C4 implementer interview, June 2015]

Collaboration, Competition, and Codesign

Codesign methodologies and intensive consumer engagement were successfully utilized in the program at scale to address knowledge gaps in consumer preferences. This work reinforced the view that a one size fits all approach would not work. However, the time and effort required for this created real challenges, as it introduced delays and consequently reduced the time available to develop and deploy new solutions within the defined timescale of the program.

 Collaboration versus competition” were also key themes. Some private industries were understandably very protective of their intellectual property, which made them unwilling to share expertise and technology solutions with third parties in a large multi-stakeholder environment.

> *...It's a sort of codesign, and what happens is people take a long time to make up their mind which compresses time for technical partners (...) for some people working in a collaborative nature with technical partners is a new environment so they are cautious and wary of telling all their secrets in case people run away with them so I think there is a protective defensive mode.* [C1 (Midpoint), e-HIT interview]

Health Service Readiness or Information Technology (IT) Infrastructure

The variation in workflow processes and in-house IT and data management systems—which stem from the historical foundations of hospitals operating as separate entities within a confederated health system—continues to impede the advancement and integration of digital health initiatives.

Technical readiness" was often an acute issue at local organizational level as legacy systems, firewalls, and strict information security procedures within health boards, hospital trusts, health centers, or general practitioner (GP) practices varied from site to site and left health professionals and implementers ill-equipped to readily deploy solutions implemented elsewhere. This became apparent as digital platforms designed and developed in a specific context could not be rolled out elsewhere at scale due to lack of process and technical standardization across the UK health services.

> *Well things like legacy systems, fire walls, when we are adopting new technologies, eHealth capacity, eHealth priorities within the internal infrastructure is stretched.* [C1, Final e-Hit interview 1]

Systems interoperability and lack of integration was a recurrent issue as computer systems across the health and social care sector could not easily exchange data with applications created as part of dallas. Practically this meant that each new implementation site had to go through its own deployment pain-barrier," with ad-hoc local solutions providing little insight on what to expect elsewhere. This limited the effectiveness of many of the new digital platforms and required workarounds by health care staff and end users to ensure the benefits could still be exploited locally.

Discontinuity and Organizational Culture

The restructuring of the health service in England meant it was particularly difficult to get health service partners to maintain focus and be ready to deliver aspects of the program they had originally signed up to as there were so many changes taking placing concurrently with the roll out of the dallas program. Constant change made ownership and responsibility for digital health services unclear and a lack of senior management buy-in was also cited as a barrier to organizational readiness for digital health, which could negatively affect implementation efforts. This recurring flux and uncertainty had a knock-on effect in other areas. Some health organizations had not yet fully developed their own digital strategies and/or had not yet bought in fully to the self-care agenda at the time of the program.

> *It's people who have to implement that need to move in a different direction to achieve what we need to achieve...It's such a complex system that you can't simply commission eHealth technology, it can't be done! There are too many stakeholders who could block, misunderstand or not want to get involved...* [C2 (Midpoint) e-HIT interview 5]

It became clear that in order to be successful, digital health innovation must be closely aligned with health service organizational vision and road maps for change. Problems at the executive level within larger health trusts were contrasted with readiness of smaller, more flexible organizations such as general practices, which seemed more receptive to adopting digital health services to large patient groups.

> *...I think (funder) should consider that on business-led projects they have to be business led, and NHS partners have to really want to do it…you only want NHS partners who see what the project is doing as something that they want to do, so the project is giving them tools or giving them insight. It's not paying them to dabble...* [C3 (Midpoint) e-HIT interview 31]

Naturally, the efforts required to implement and manage changes are also less complex and costly for smaller organizations and this program allowed them to take risks and test out digital health at scale. Thus, smaller organizations such as consortia of primary care practices were able to be more flexible and more responsive to opportunities presented by dallas communities.

Resource Constraints

The health service and other organizations frequently had to contend with major resource constraints during the current period of financial austerity, which affected ability to engage with various initiatives within dallas.

> *...we're quite far behind in our IT. We don't have electronic records as such, so we're still writing in records. And I think that's probably half the problem.* [Health visitor, Focus Group, 2015]

Budget constraints were clearly visible among some partners who struggled with manpower capacity and to provide mobile technology and other equipment for their health care staff to roll out different digital health products and services, which was a major impediment to progressing the digital health consumer" agenda.

Micro-Level Readiness

Two overarching themes were identified at the micro level: readiness of health professionals" and readiness of the public and patients."

Health Professional Readiness: Workload and Professional Confidence

Workload pressures and lack of capacity was a recurring barrier to incorporating new technologies into everyday working practices and on occasion incentivization had to be used to overcome this.

> *...it was more difficult at the start but as we have got more patients onto the system, as we have started to be able to say you know this is what the patients are saying, GP practices are warmed to what we are doing and actually become proactive themselves in trying to get their patients onto telehealth. We did have a £600 payment that we would give the GP practices for supplying us with their long term conditions list (...) so it's just a locally enhanced scheme...Now to be honest they don't even talk about the money now.* [C2, Final e-HIT interview 2, health service manager]

Concerns around data security, when partnering with private companies, were a barrier to uptake. Some nurses worried that personal information may not be kept confidential and secure or could be sold or shared with commercial companies without the explicit consent or knowledge of patients. Also, given that many digital health initiatives have come and gone over the years and failed to be integrated into routine care in the UK health service, some health professionals were skeptical about the likely longevity of the dallas program and were slow to engage with the technologies it was promoting. In these cases, change fatigue" was evident.

Training and Alignment With Professional Roles and Identities

There was a perception by those tasked with driving forward new initiatives that clinicians and others feared digital health as they believed it could be used to disempower and in some cases replace them as care providers.

> *that's quite a big initial thought of a lot of different care groups…, that they'd be made redundant by the introduction of technology"* [C2 e-HIT interview, representative charity organization]

However, the main barrier was not a fear of role replacement, but lack of knowledge and skills in relation to digital health which significantly affected how prepared some were to engage with the different technologies.

> *...I think the whole system about IT, I am a nurse and that's what I was trained to do, so before IT came in, we were doing everything on paper, and now things are changing for us, and we've never really been given training, we're only doing it on the job, and we've had a new IT system called "x" coming in, that's created an absolute nightmare for everybody, because we're not necessarily that skilled in IT processes to be able to do that. So generalized IT training would be good...then tailoring it to the things that we're doing...* [Health visitor, Focus group, April 2015]

The pace of technological change was also noted as a problem as clinicians felt unable to keep up to date with new and emerging developments, especially the growth and lack of regulation within the mobile app market.

Access to Digital Resources

Even when staff were digitally literate, some found it impossible to drive new digital health services forward because of the technical infrastructure issues reported previously. For example, poor connectivity of mobile devices such as tablet computers and firewalls that blocked their access to internet and social media on NHS sites prevented engagement with new digital health applications and services.

> *...the key issues were about access to equipment...it was the statutory sector that struggled. ...the security systems that we have on most of our desktops actually stop you getting access to things like (new dallas digital tool). So the browsers were so old that a lot of the material wasn't displaying properly when you were sitting on the public sector end…so it was more to do with the challenges roundabout making sure that staff in the public sector had access to the level of equipment that people take for granted in their day-to-day lives.* [C1, implementer interview, June 2015]

Public Readiness: Digital Literacy and Access

Variation in digital literacy skills caused widespread problems across dallas. The target market ranged from younger consumers—who were more adept, confident, and ready to use digital tools—through to some older adults with little or no previous awareness or understanding of basic IT.

> *...quite a few of them had no digital knowledge whatsoever, they had no access, they had nobody that was able to show people how to use digital stuff...* [Digital champion interview, March 2015]

Some consortia had specifically undertaken community-asset based activities to address digital access and inclusion such as setting up digital hubs and creating digital champions to encourage people to get on the Web as many individuals were clearly less digitally "able." There appeared to be a risk of compounding the inverse care law," with those from lower socio-economic groups—often most in need of health and social care services—being less able to digitally" access these.

> *...they give us these and said "Here you are! Here are the tablets," the first we had seen a tablet was: you took it with water and you put it in your mouth."* [House of Memories," Focus group, March 2015]

Despite reported growing use of digital tablets and smartphones among the general population, many people still lacked basic access to such mobile devices. The cost of technology or poor access to computer equipment and free Internet services in local communities prevented many individuals from participating in some of the dallas offerings. To overcome this, one consortium—serving a mixed population including communities in high deprivation—actively sought to develop new routes of access, such as providing education and digital support, as previously reported.

Agency of Individuals and Their Perceptions of Consumer” Digital Health Tools

Technologies which required data entry and/or a change in daily routines affected the readiness” of users to adopt these and this proved another challenge for the preventive self-care agenda. The issue of individual agency” arose as people lead busy lives with competing career and family responsibilities, as well as—for some—financial and social demands, which were often prioritized ahead of health.

> *…People don’t prioritize health, so if you are economically deprived, you prioritize feelings of physical safety and financial safety so you could be worried about paying your rent, keeping debt collectors off the door, anti-social behavior in your neighborhood. If you’re more economically active then other things are a priority, holidays, kids, schooling, housing, mortgage.* [C2 implementer interview, June 2015]

Some health professionals expressed concern that some mobile interventions were not always necessarily appropriate for their service users—both older and younger—depending upon timing and/ or settings. Equally, while some welcomed new digital solutions, others did not. For example, some individuals were eager to access digital health solutions for elderly parents only to find that the parents themselves were extremely reluctant to permit some form of home monitoring.”

In addition, the increasing multicultural nature of contemporary society presented a barrier when English was the only language available on the digital platforms.

> *...The only people that I can—hand on heart—say I haven’t offered it to since we started to do this have been a couple of my Polish clients that haven’t spoken English, the vast majority of them speak enough and if they speak and if they can understand me in the booking I will say would you like to access this and often if it’s spoken they can...I have had a couple that have come and there is just not a word of English and for that reason I have chosen not to go there.* [Community midwife, Focus group, April 2015]

Trust in Consumer-facing” Digital Technologies

Trust in digital health security was a persistent issue with some expressing unease about the safety and security of privately held data and whether or not it would be shared with organizations without their explicit knowledge or consent, given recent high profile data breaches.

> *...I think perception of risk to patient data is a big challenge. People are uncertain about the implications of sharing their data with a system and well it’s difficult to explain the subtleties of the distinction between personally held record which they own the data…while we understand the concept of anonymizing data and who owns consent and everything else those are quite complicated messages to pass to the general public.* [C4 final e-HIT interview 4, industry representative]

Clinical endorsement and validation was seen as one way to address this and promote uptake and utilization among consumer groups.

Discussion

Our results show that readiness issues have been ubiquitous across macro, meso, and micro levels and across sectoral boundaries: market, policy, organizational, professional, and consumer. These issues are not insurmountable challenges but their existence does need to be acknowledged and addressed if deployment at scale to the widest population is to be realized.

Study Strengths and Limitations

We have examined the implementation journey of a national program aiming to deliver digital health at scale across the United Kingdom. We have rigorously collected and analyzed process data from a wide range of stakeholders involved in the implementation, identifying not only potential barriers, but also why these occur and how to address them in future (Table 4). A robust, sociological theory—NPT—underpinned the evaluation as recommended as good practice [32,33]. This could result in inappropriate shoe horning” of data, however, we would argue that we were rigorous in looking for data that also fell outside the framework. Our qualitative data collection was largely limited to those engaging with the dallas program and we did not undertake work with individuals or organizations that were unable or unwilling to engage with the program, which could have provided different perspectives and possibly shed light on digital exclusion or nonparticipation.

Our research was located in 2 countries within the United Kingdom which operate a NHS system where health care is free at the point of access and there are funding constraints which is important to note when considering the implications of this work. In addition, the governments in Scotland and England have a major role in developing and overseeing regulatory and information governance frameworks. Finally, there is a long history of embracing digitalization in health care, for example, all primary care practitioners use electronic medical record systems, which is important to note when comparing the United Kingdom with less digitally advanced countries.

How Does This Study Fit With the Existing Literature?

Consumer adoption of digital health is seen as a great market opportunity with numerous policy drivers and yet penetration of this large potential market remains relatively poor [37]. Our findings resonate with reports and data from other sources, for example, digital skill and infrastructure deficiencies have been noted by a recent select committee report on digital skills for 2014-2015 [38] that examined challenges for a digital economy” more broadly. The UK government has recognized this issue and vowed to make fast broadband available to every home [39] but our findings suggest this will still leave much to do. Interoperability is a key aim in the United Kingdom as it strives to implement digital standards and achieve system wide interoperability but others have reported interoperability as a barrier to implementation and large scale deployments of mHealth and global health interventions [40,41]. The recent European Union (EU) Task Force on eHealth also acknowledged the need to develop EU-wide standards on interoperability and data sharing [26]. Our whole system view of the digital health ecosystem provides potential explanations, suggesting that interoperability is not a technical issue but rather due to industry inertia and to multiple organizations operating within the health system in the United Kingdom.

Clinical endorsement of digital health products and services including systems for regulation and accreditation of technology and data enabled services is required and has been suggested previously [42], and a recent systematic review of the international literature on barriers and facilitators to patient and public involvement with digital health has suggested this is a key issue [43]. Our work suggests that the health care community would welcome better integration of health records, although persistent challenges are posed by the way current information governance rules are interpreted and enacted. These types of problems need to be addressed if the aim is to share data across sectors. Previous research relating to the use of personal electronic health records has demonstrated less public and professional appetite than anticipated [22,44].

Our work suggests that problems identified in the wider literature on diffusion of innovations such as the importance of structural determinants (such as resources), definite perceived advantages of the innovation, ease of use, good training and support, as well as ability to address perceived risks of new ways [3] of working apply equally to the digital health sphere. Importantly the dallas program has shown that although such issues persist across a range of digital health initiatives more user centered design techniques, intensive engagement, and support of users and incentivization of professionals can help increase interest in digital health.

Implications for Implementations of Digital Health Technologies at Scale

The dallas program highlighted challenges but also potential solutions to the large scale implementation of digital health, for example, through the development of information governance recommendations for health care organizations [45] and the use of digital champions to address skill deficiencies. Our findings lead us to a set of actionable recommendations for future work and for increasing readiness for digital health at scale (Table 4).

Table 4. Recommendations for future implementation work in digital health.

Recommendation no.	Recommendation
Recommendation 1	Further commitment and investment in both national and local infrastructure will be required if digital health care is to become normalized.
Recommendation 2	Guidance relating to ownership and control of personal health data and data privacy regulations are required to mitigate current uncertainty in the digital health arena.
Recommendation 3	Brand trust and confidence is crucial. Accreditation and official endorsement of products and services is an important determinant of future successful deployment of digital health services as is peer recommendation for consumer wellness products. Clear systems to facilitate trust and confidence need to be put in place.
Recommendation 4	Technical and service interoperability needs to be prioritized and, if necessary, incentivized to ensure the scaling up of digital health care across systems and sectors.
Recommendation 5	Future digital health services need to be more accessible by those who are currently socially or economically excluded including those whose first language is not English, and those with sensory, physical, or cognitive impairments.
Recommendation 6	There is a need to invest in further awareness raising, upskilling of consumers and more affordable and accessible technologies if the true potential of digital health and wellbeing technologies are to be fully realized and the concept of professional and lay champions to promote technologies and services merit support.
Recommendation 7	More extensive and intensive public engagement and debate on the subject of the risks versus benefits of digital health needs to be undertaken to address concerns around security and safety of digital health and wellness products and services.
Recommendation 8	Greater emphasis needs to be placed on both upskilling and also ensuring the next generation of health professionals are more digitally” able. Digital health care needs to be a feature of undergraduate health professional training.
Recommendation 9	Guidance is required to shape and support a market that spans consumer wellness and statutory health services. Consideration must be given to future funding models, procurement, and the potential for hybrid data, including sharing, storage, and management models that permit digital health apps and services to be taken up and used via consumer markets and/or statutory channels.
Recommendation 10	There is a need to promote health care stability and a culture of long term planning. Instability and constant change can be a deterrent to investment and hinders implementation in the digital health sphere.

Conclusions

Although there is much rhetoric about the consumer push for digital health, our research raises some outstanding issues relating to the readiness for digital health that need attention. We provide a set of 10 key recommendations that aim to tackle these issues. If addressed, these recommendations will promote the right market and environment to permit the routine deployment and true scaling-up of digital health and wellness technologies and services.

Acknowledgments

The authors wish to express thanks to the funding body Innovate UK, the 4 dallas communities and the dallas evaluation advisory board which includes Anne Rogers, Alan Connell, Carl May, Christine Asbury, David Jarrold, Jeremy Linksell, Joy Todd, Mike Biddle, Nick Goodwin, Stanton Newman, Tim Ellis, and William Maton-Howarth. The dallas programme was funded by Innovate UK (formerly known as the U.K. Technology Strategy Board). The views expressed in this paper are those of the authors and not necessarily those of Innovate UK or any of the dallas funding bodies.

Authors' Contributions

FSM, MML, MMB, COD, EG, NW, SW, and TF conceived and designed the dallas evaluation framework and secured funding for this work. MMB, MML, RA, SOC, AD, and AB collected the dallas qualitative dataset. AMD, MML, SOC, UC, RA, COD, and FSM analyzed the dataset. FSM, MML, MMB, AMD, and SOC drafted the manuscript with all coauthors helping to refine the draft and revise it. Lead and corresponding authors (MML and FM) both affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. All authors approved the final manuscript.

Conflicts of Interest

All authors have declared that they (1) have support from University of Glasgow, University of Strathclyde, University of Newcastle and Innovate UK for the submitted work, (2) have no relationships with any companies that have an interest in the submitted work in the previous 3 years, (3) their spouses, partners, or children have no financial relationships relevant to the submitted work, and (4) have no other nonfinancial interests that may be relevant to the submitted work.

Multimedia Appendix 1

Overview of dallas communities.

[PDF File (Adobe PDF File), 103KB - jmir_v19i2e42_app1.pdf]

Multimedia Appendix 2

The 4 normalization process theory (NTP) constructs.

[PDF File (Adobe PDF File), 36KB - jmir_v19i2e42_app2.pdf]

Multimedia Appendix 3

Qualitative tables with additional study participants' quotes.

[PDF File (Adobe PDF File), 98KB - jmir_v19i2e42_app3.pdf]

References

1. Jennett PA, Andruchuk K. Telehealth: 'real life' implementation issues. Comput Methods Programs Biomed 2001 Mar;64(3):169-174. [Medline: 11226614]
2. 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]
3. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629 [FREE Full text] [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
4. Mair F, 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 01;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
5. Bowes A, McColgan G. Implementing telecare for people with dementia: supporting ageing in place in West Lothian, Scotland. J Care Serv Manag 2013 Jul 18;3(3):227-243. [doi: 10.1179/csm.2009.3.3.227]

6. Hibbert D, Mair F, Angus R, May C, Boland A, Haycox A, et al. Getting beyond “trial and error”: Lessons from the implementation of a home telecare service. J Telemed Telecare 2003;9(1):55-56.
7. May CR, Finch TL, Cornford J, Exley C, Gately C, Kirk S, et al. Integrating telecare for chronic disease management in the community: what needs to be done? BMC Health Serv Res 2011 May 27;11:131 [FREE Full text] [doi: 10.1186/1472-6963-11-131] [Medline: 21619596]
8. Kruse C, Regiekr V, Rheinboldt K. Barriers over time to full implementation of health information exchange in the United States. JMIR Med Inform 2014;2(2):e26. [doi: 10.2196/medinform.3625]
9. Finch TL, Mair FS, May CR. Teledermatology in the UK: lessons in service innovation. Br J Dermatol 2007 Mar;156(3):521-527. [doi: 10.1111/j.1365-2133.2006.07608.x] [Medline: 17300243]
10. Goroll AH, Simon SR, Tripathi M, Ascenzo C, Bates DW. Community-wide implementation of health information technology: the Massachusetts eHealth collaborative experience. J Am Med Inform Assoc 2009;16(1):132-139 [FREE Full text] [doi: 10.1197/jamia.M2899] [Medline: 18952937]
11. Sheikh A, Cornford T, Barber N, Avery A, Takian A, Lichtner V, et al. Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from prospective national evaluation in “early adopter” hospitals. Br Med J 2011 Oct 17;343:1-14. [doi: 10.1136/bmj.d6054]
12. Jennett P, Jackson A, Ho K, Healy T, Kazanjian A, Woollard R, et al. The essence of telehealth readiness in rural communities: an organizational perspective. Telemed J E Health 2005 Apr;11(2):137-145. [doi: 10.1089/tmj.2005.11.137] [Medline: 15857254]
13. Nahm ES, Blum K, Scharf B, Friedmann E, Thomas S, Jones D, et al. Exploration of patients' readiness for an eHealth management program for chronic heart failure: a preliminary study. J Cardiovasc Nurs 2008;23(6):463-471. [doi: 10.1097/01.JCN.0000338930.89210.79] [Medline: 18953206]
14. Khoja S, Scott RE, Casebeer AL, Mohsin M, Ishaq AF, Gilani S. e-Health readiness assessment tools for healthcare institutions in developing countries. Telemed J E Health 2007 Aug;13(4):425-431. [doi: 10.1089/tmj.2006.0064] [Medline: 17848110]
15. Finch TL, Mair FS, O'Donnell C, Murray E, May CR. From theory to 'measurement' in complex interventions: methodological lessons from the development of an e-health normalisation instrument. BMC Med Res Methodol 2012 May 17;12:69 [FREE Full text] [doi: 10.1186/1471-2288-12-69] [Medline: 22594537]
16. Koopman RJ, Petroski GF, Canfield SM, Stuppy JA, Mehr DR. Development of the PRE-HIT instrument: patient readiness to engage in health information technology. BMC Fam Pract 2014 Jan 28;15:18 [FREE Full text] [doi: 10.1186/1471-2296-15-18] [Medline: 24472182]
17. May C, Mort M, Williams T, Mair F, Gask L. Health technology assessment in its local contexts: studies of telehealthcare. Soc Sci Med 2003 Aug;57(4):697-710. [Medline: 12821017]
18. Jennett PA, Scott RE, Affleck HL, Hailey D, Ohinmaa A, Anderson C, et al. Policy implications associated with the socioeconomic and health system impact of telehealth: a case study from Canada. Telemed J E Health 2004;10(1):77-83. [doi: 10.1089/153056204773644616] [Medline: 15104919]
19. Childs S, Blenkinsopp E, Hall A, Walton G. Effective e-learning for health professionals and students--barriers and their solutions. A systematic review of the literature--findings from the HeXL project. Health Info Libr J 2005 Dec;22(Suppl 2):20-32 [FREE Full text] [doi: 10.1111/j.1470-3327.2005.00614.x] [Medline: 16279973]
20. Lehmann ED. Computerised decision-support tools in diabetes care: hurdles to implementation. Diabetes Technol Ther 2004 Jun;6(3):422-429. [doi: 10.1089/152091504774198142] [Medline: 15198848]
21. Greenhalgh T, Swinglehurst D, Stones R. Rethinking resistance to 'big IT': a sociological study of why and when healthcare staff do not use nationally mandated information and communication technologies. Health Services and Delivery Research 2014;2(39):-. [Medline: 27466649]
22. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. Br Med J 2010 Nov 16;341:c5814 [FREE Full text] [Medline: 21081595]
23. Devlin AM, McGee-Lennon M, O'Donnell CA, Bouamrane MM, Agbakoba R, O'Connor S, et al. Delivering digital health and well-being at scale: lessons learned during the implementation of the dallas program in the United Kingdom. J Am Med Inform Assoc 2016 Jan;23(1):48-59 [FREE Full text] [doi: 10.1093/jamia/ocv097] [Medline: 26254480]
24. Brown MJ. Personalised medicine for hypertension. Br Med J 2011 Jul 28;343:d4697. [Medline: 21798973]
25. Agusti A. The path to personalised medicine in COPD. Thorax 2014 Sep;69(9):857-864. [doi: 10.1136/thoraxjnl-2014-205507] [Medline: 24781218]
26. Ec.europa. Luxembourg: EU Task Force on eHealth; 2014. EU Task Force on eHealth: Redesigning health in Europe for 2020 URL: https://ec.europa.eu/digital-single-market/en/news/eu-task-force-ehealth-redesigning-health-europe-2020 [accessed 2017-02-03] [WebCite Cache ID 6o0IwAk4E]
27. Gov.UK. Personalised Health and Care 2020: Using data and technology to transform outcomes for patients and citizens URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384650/NIB_Report.pdf [accessed 2017-02-03] [WebCite Cache ID 6o0HUjm5x]
28. Murray E, May C, Mair F. Development and formative evaluation of the e-Health Implementation Toolkit (e-HIT). BMC Med Inform Decis Mak 2010 Oct 18;10:61 [FREE Full text] [doi: 10.1186/1472-6947-10-61] [Medline: 20955594]

29. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015:2015. [doi: 10.1186/s13012-015-0242-0]
30. Cresswell KM, Worth A, Sheikh A. Actor-Network Theory and its role in understanding the implementation of information technology developments in healthcare. BMC Med Inform Decis Mak 2010 Nov 01;10:67 [FREE Full text] [doi: 10.1186/1472-6947-10-67] [Medline: 21040575]
31. Weiner BJ. A theory of organizational readiness for change. Implement Sci 2009 Oct 19;4:67 [FREE Full text] [doi: 10.1186/1748-5908-4-67] [Medline: 19840381]
32. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. Sociology 2009 Jun;43(3):535-554. [doi: 10.1177/0038038509103208]
33. McEvoy R, Ballini L, Maltoni S, O'Donnell C, Mair FS, MacFarlane A. A qualitative systematic review of studies using the normalization process theory to research implementation processes. Implement Sci 2014;9:1-13. [doi: 10.1186/1748-5908-9-2]
34. Finch T, Girling M, May C, Mair F, Murray E, Treweek S, et al. Normalizationprocess. 2015. NoMAD: Implementation measure based on Normalization Process Theory URL: http://www.normalizationprocess.org/ [accessed 2017-02-03] [WebCite Cache ID 6o0HsGWPO]
35. Miles M. Qualitative Data Analysis for Applied Policy Research, in The Qualitative Researchers Companion, A.M. Huberman, MB. In: The qualitative researcher's companion. Thousand Oaks, CA: Sage Publications; 2002.
36. Health D. Gov.UK. 2013. Caldicott Review: information governance in health and care system, Information: To share or not to share? URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf [accessed 2017-02-03] [WebCite Cache ID 6o0IB1Ana]
37. Gov.UK. 2015 Sep. Digital Health in the UK: An industry study for the Office of Life Sciences URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/461479/BIS-15-544-digital-health-in-the-uk-an-industry-study-for-the-Office-of-Life-Sciences.pdf [accessed 2017-02-03] [WebCite Cache ID 6o0IJHENi]
38. Publications.parliament. 2015. Make or Break the UKs Digital Future URL: https://www.publications.parliament.uk/pa/ld201415/ldselect/lddigital/111/111.pdf [accessed 2017-02-03] [WebCite Cache ID 6o0IRgnP8]
39. Department of Culture. Gov.UK. 2013. Broadband Delivery UK URL: https://www.gov.uk/guidance/broadband-delivery-uk [accessed 2017-02-03] [WebCite Cache ID 6o0IaFGKS]
40. Kostkova P. Grand challenges in digital health. Front Public Health 2015;3:134 [FREE Full text] [doi: 10.3389/fpubh.2015.00134] [Medline: 26000272]
41. Bouamrane M, Tao C, Sarkar IN. Managing interoperability and complexity in health systems. Methods Inf Med 2015;54(1):1-4. [doi: 10.3414/ME15-10-0001] [Medline: 25579862]
42. Huckvale K, Prieto J, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015 Sep 07;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
43. 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]
44. Serrano K, Yu M, Riley W, Patel V, Hughes P, Marchesini K, et al. Willingness to exchange health information via mobile devices: findings from a population-based survey. Ann Fam Med 2016;14(1):34-40 [FREE Full text] [doi: 10.1370/afm.1888] [Medline: 26755781]
45. Chitty A. HSJ.: Health Services Journal; 2015 Jun 22. Make people the controllers of their data to help the NHS go digital URL: http://www.hsj.co.uk/5086634.article#.VYkov2jF99d [accessed 2017-02-03] [WebCite Cache ID 6o0GnQ08Q]

Abbreviations

dallas: Delivering Assisted Living Lifestyles at Scale
NPT: normalization process theory

Edited by G Eysenbach; submitted 27.10.16; peer-reviewed by C Auschra, S Clough, I Montagni; comments to author 17.11.16; revised version received 08.12.16; accepted 13.01.17; published 16.02.17.

Please cite as:
Lennon MR, Bouamrane MM, Devlin AM, O'Connor S, O'Donnell C, Chetty U, Agbakoba R, Bikker A, Grieve E, Finch T, Watson N, Wyke S, Mair FS
Readiness for Delivering Digital Health at Scale: Lessons From a Longitudinal Qualitative Evaluation of a National Digital Health Innovation Program in the United Kingdom
J Med Internet Res 2017;19(2):e42
URL: http://www.jmir.org/2017/2/e42/
doi:10.2196/jmir.6900
PMID:28209558

Original Paper

Virtual Visits for Acute, Nonurgent Care: A Claims Analysis of Episode-Level Utilization

Aliza S Gordon[1], MPH; Wallace C Adamson[2], MD; Andrea R DeVries[1], PhD

[1]HealthCore, Inc, Wilmington, DE, United States
[2]Anthem, Inc, Indianapolis, IN, United States

Corresponding Author:
Aliza S Gordon, MPH
HealthCore, Inc
123 Justison Street
Suite 200
Wilmington, DE, 19801
United States
Phone: 1 302 230 2007
Fax: 1 302 230 2020
Email: agordon@healthcore.com

Abstract

Background: Expansion of virtual health care—real-time video consultation with a physician via the Internet—will continue as use of mobile devices and patient demand for immediate, convenient access to care grow.

Objective: The objective of the study is to analyze the care provided and the cost of virtual visits over a 3-week episode compared with in-person visits to retail health clinics (RHC), urgent care centers (UCC), emergency departments (ED), or primary care physicians (PCP) for acute, nonurgent conditions.

Methods: A cross-sectional, retrospective analysis of claims from a large commercial health insurer was performed to compare care and cost of patients receiving care via virtual visits for a condition of interest (sinusitis, upper respiratory infection, urinary tract infection, conjunctivitis, bronchitis, pharyngitis, influenza, cough, dermatitis, digestive symptom, or ear pain) matched to those receiving care for similar conditions in other settings. An episode was defined as the index visit plus 3 weeks following. Patients were children and adults younger than 65 years of age without serious chronic conditions. Visits were classified according to the setting where the visit occurred. Care provided was assessed by follow-up outpatient visits, ED visits, or hospitalizations; laboratory tests or imaging performed; and antibiotic use after the initial visit. Episode costs included the cost of the initial visit, subsequent medical care, and pharmacy.

Results: A total of 59,945 visits were included in the analysis (4635 virtual visits and 55,310 nonvirtual visits). Virtual visit episodes had similar follow-up outpatient visit rates (28.09%) as PCP (28.10%, *P*=.99) and RHC visits (28.59%, *P*=.51). During the episode, lab rates for virtual visits (12.56%) were lower than in-person locations (RHC: 36.79%, *P*<.001; UCC: 39.01%, *P*<.001; ED: 53.15%, *P*<.001; PCP: 37.40%, *P*<.001), and imaging rates for virtual visits (6.62%) were typically lower than in-person locations (RHC: 5.97%, *P*=.11; UCC: 8.77%, *P*<.001; ED: 43.06%, *P*<.001; PCP: 11.26%, *P*<.001). RHC, UCC, ED, and PCP were estimated to be $36, $153, $1735, and $162 more expensive than virtual visit episodes, respectively, including medical and pharmacy costs.

Conclusions: Virtual care appears to be a low-cost alternative to care administered in other settings with lower testing rates. The similar follow-up rate suggests adequate clinical resolution and that patients are not using virtual visits as a first step before seeking in-person care.

(J Med Internet Res 2017;19(2):e35) doi:10.2196/jmir.6783

KEYWORDS

virtual visit; health care utilization; claims analysis

Introduction

Health care delivery is moving outside traditional settings of physician offices and emergency departments (EDs) into convenient quick-care sites [1] such as retail health clinics (RHCs) and urgent care centers (UCCs). In addition to these relatively established alternatives for nonurgent acute care [2,3], venues maximizing the latest technology are emerging. Telehealth includes both asynchronous structured e-visits and online synchronous live video visits (virtual visits). Structured e-visits refer to online communication (eg, inputting symptoms into a website) and provision of treatment plans over the Internet, while virtual visits feature real-time video consultation. By providing a means for patients to receive health care from any location at all hours, telehealth platforms expand patient access to medical care [1,4,5]. While the American College of Physicians supports telemedicine, its policy statement stresses the importance of maintaining the physician-patient relationship [6].

Because telehealth and particularly virtual visits are relatively new care options, published literature is lacking, with previous studies largely focused on acceptability or patient characteristics rather than outcomes [7-11]. Quality issues are particularly concerning for telehealth, such as whether physicians can provide accurate diagnoses without hands-on physical examinations of patients, whether patients receive appropriate laboratory testing after the visit, and whether antibiotics are overprescribed [12-14]. There are also concerns about whether the visits are truly efficient—if additional follow-up visits occur as a result of unresolved symptoms or because patients used online visits as a first opinion before seeking care in an office setting, this would diminish the apparent cost-saving potential. The few studies that examined the quality of virtual visits found higher antibiotic use after virtual visits compared with in-person office visits [13]—including prescription of broad-spectrum antibiotics [14]—for acute respiratory infections where antibiotics are not generally recommended. Providers of virtual visits were less likely than those in offices to order diagnostic tests to determine whether cases of pharyngitis were bacterial or viral, which is considered the standard of care [13]. Using follow-up visits as a proxy measure of misdiagnoses or treatment failures, researchers found a similar rate among patients participating in structured e-visits as those going to an office [6,12]. While an initial telehealth visit (e-visit or virtual) is usually less costly than an in-person visit, few if any studies analyzed episode-level costs associated with virtual visits (including follow-up care), although 1 study of structured e-visits found treatment costs per episode of care to be lower than traditional settings [7].

Virtual visits have recently become available through independent online health care delivery sites that are often covered by patients' health plans [14]. Starting in 2014, virtual visits became available to the members in this study as a covered benefit, with patient copays similar to primary care physicians (PCP) office visits. After patients create an initial profile, they choose from a list of available providers licensed in their state of residence [15]. Physicians are reimbursed $49 per virtual visit. The service has become increasingly popular with members, who received up to 4000 virtual visits per month through this service in 2015. Patient satisfaction has been high, with a net promoter score of 65%, based on an exit survey of patients who used the service administered as part of the health plan's virtual care program (personal communication, W Adamson).

To expand our understanding of the care provided and costs associated with virtual health care, we examined the care for specific acute, nonurgent conditions (eg, colds, allergies, urinary tract infections) provided by physicians via a virtual visit platform. Care provided through virtual visits, including subsequent care during a 3-week follow-up period, was compared with care delivered in the RHC, UCC, ED, and PCP office settings. This study is unique in assessing care and costs of virtual visit episodes, in contrast to previous studies of telehealth costs that have assessed structured e-visits only or have not taken into account follow-up care after the initial visit.

Methods

Study Design

This cross-sectional retrospective study used data from commercially insured members receiving virtual care matched to members receiving care for similar conditions in other settings. The claims-based dataset was derived from the HealthCore Integrated Research Database (HIRD), a large administrative claims database containing medical and pharmacy claims for 14 Anthem commercial health plans geographically dispersed across the United States. The patient sample was identified from claims with service dates during the study period, January 1, 2014, through May 11, 2015. Researchers had access to a limited dataset containing no patient identifiers. This study was conducted in full compliance with the Health Insurance Portability and Accountability Act. This study was nonexperimental and was exempt from investigational review board approval.

The index date was defined as the date of the first outpatient or ED claim in a 3-week period for 11 of the most commonly diagnosed conditions through the telehealth platform: sinusitis, upper respiratory infections (URIs), urinary tract infections (UTIs), conjunctivitis, bronchitis, pharyngitis, influenza, cough, dermatitis, nausea/vomiting/diarrhea, and ear pain, based on *International Classification of Diseases, Ninth Revision* (ICD-9) diagnosis codes (Multimedia Appendix 1). Baseline patient characteristics were determined from claims during the 6 months prior to the index date.

Selection Criteria

The study included adults younger than 65 years of age and children who had health plan eligibility for at least 6 months before and 3 weeks after the index date. Patients with serious or expensive health conditions, defined by Deyo-Charlson Comorbidity Index (DCI) scores of greater than 2, or with cystic fibrosis, transplant, end-stage renal disease, HIV, hemophilia, stroke, or respiratory failure were excluded.

Episode Identification

Visits for conditions of interest were classified according to the setting where the visit occurred: virtual (identified by Current Procedural Terminology [CPT] code 99444 and tax ID, representing all covered telehealth visits), RHCs (identified by tax ID and National Provider Identifier [NPI] numbers), UCCs (identified by tax ID and NPI numbers; only large national UCC chains included), EDs (identified by revenue codes and CPT codes), and PCP offices (identified by CPT codes for outpatient evaluation and management visits with provider specialty noted as primary care, internal medicine, general medicine, or pediatrics).

An episode was defined as the index visit plus 3 weeks following. If patients had 2 or more potential index visits less than 3 weeks apart, only the first visit was used to identify an episode. If patients had more than 1 visit on the same day, a hierarchy was used to determine the index visit (as opposed to follow-up visit). The hierarchy was virtual > RHC > PCP > UCC > ED; this order was chosen due to likelihood of patients going to a more "urgent" care location after a different option was tried if more than 1 location was visited in a single day.

Patient Selection

Members receiving care from RHC, UCC, ED, and PCP offices were matched to those with virtual visits in a 3:1 ratio for each location (to increase statistical power) on acute condition, quarter and year of index date, state/region of residence, and child (<18 years) or adult age group.

Care and Cost Outcome Measures

The primary outcome measures were care provided (utilization during and following the visit) and cost of care. The follow-up period for outcomes assessment was from the index date to 3 weeks after to allow sufficient time for most minor conditions to resolve [16]. The follow-up period for antibiotic fills was 3 days from the index date since most antibiotic fills occurred during this time and fills occurring later in the episode may have occurred after a follow-up visit at a different location.

Care provided was assessed by subsequent medical care after the initial visit (ie, outpatient evaluation and management visit [follow-up visit], ED visit, or inpatient hospitalization), laboratory tests performed, imaging performed, and antibiotic fill rates and use of broad-spectrum antibiotics, for patients where pharmacy data were available. Allowed cost per episode included the cost of the initial visit, subsequent medical care, and pharmacy costs.

All care and costs during the 3-week episode were included, not just those for care with the same diagnosis as the index visit, since it is difficult to determine whether subsequent care is related to the initial visit (eg, pneumonia can develop after a different infection).

Statistical Analysis

The outcome measures were analyzed to determine differences between virtual visits and other locations of care. A significance level of $\alpha<.05$ (2-sided) comparing each location with virtual was considered for all analyses (*P* values compare each location to virtual). Frequencies were reported and χ^2 tests were used for all differences in care patterns/utilization. General linear models with gamma distribution and log link were used to compare costs and were adjusted to account for differences in age category and common baseline comorbidities (see Table 2 for specific age categories and comorbidities). We did not adjust for prior costs or utilization since prior care seeking is related to future care seeking behavior and does not necessarily match with health status. Due to the large variation in and skewed nature of follow-up costs, a sensitivity analysis was performed using winsorized values at the 5% and 95% level. Statistical analyses were conducted using SAS version 9.4 (SAS Institute, Inc).

Results

Selected Visits

A total of 4635 virtual and 55,310 in-person visits were included in the analysis (13,832 RHC; 13,757 UCC; 13,840 ED; 13,881 PCP; see Figure 1 and Table 1). Pharmacy data were available for 3182 virtual and 29,562 other visits (7518 RHC; 7188 UCC; 7227 ED; 7629 PCP). The condition mix in each group consisted of the proportion of virtual visits with each of the 11 conditions (Table 1). Sinusitis and URI were the most highly represented conditions, accounting for more than half of the sample.

Figure 1. Attrition; number of virtual visits at each step.

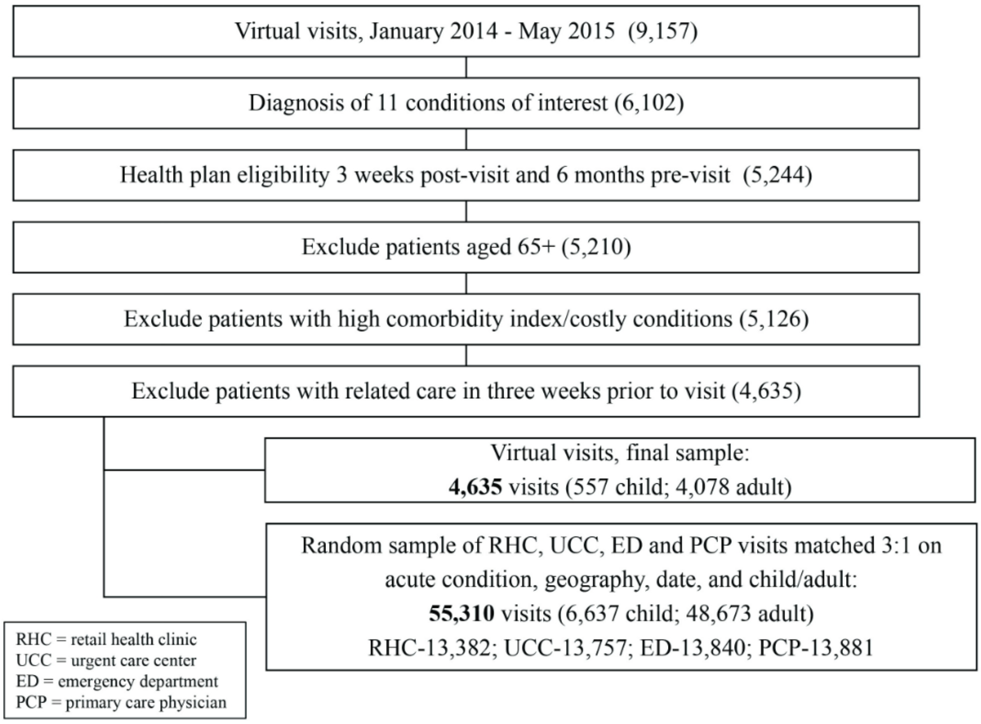


Table 1. Visits by diagnosis.

Diagnosis	Virtual, n (%)	RHC[a], n (%)	UCC[b], n (%)	ED[c], n (%)	PCP[d], n (%)
Sinusitis	1689 (36.44)	5055 (36.55)	5062 (36.80)	5029 (36.34)	5060 (36.45)
Upper respiratory infection	849 (18.32)	2540 (18.36)	2534 (18.42)	2537 (18.33)	2541 (18.31)
Urinary tract infection	413 (8.91)	1240 (8.96)	1238 (9.00)	1236 (8.93)	1239 (8.93)
Bronchitis	397 (8.57)	1191 (8.61)	1195 (8.69)	1188 (8.58)	1192 (8.59)
Conjunctivitis	356 (7.68)	1070 (7.74)	1035 (7.52)	1071 (7.74)	1068 (7.69)
Pharyngitis	285 (6.15)	854 (6.17)	853 (6.20)	851 (6.15)	853 (6.15)
Cough	158 (3.41)	471 (3.42)	473 (3.44)	469 (3.39)	472 (3.40)
Contact dermatitis	145 (3.13)	435 (3.14)	410 (2.98)	432 (3.12)	432 (3.11)
Influenza	140 (3.02)	418 (3.02)	386 (2.81)	419 (3.03)	417 (3.00)
Digestive symptoms—diarrhea, nausea, vomiting	104 (2.24)	260 (1.88)	310 (2.25)	312 (2.25)	311 (2.24)
Ear disorders—ear pain	99 (2.14)	296 (2.14)	261 (1.90)	296 (2.14)	296 (2.13)
Total	4635 (100)	13,832 (100)	13,757 (100)	13,840 (100)	13,881 (100)

[a] RHC: retail health clinic.

[b] UCC: urgent care center.

[c] ED: emergency department.

[d] PCP: primary care physician.

Baseline Patient Characteristics

In the RHC, UCC, and ED groups, the highest proportion of patients were 18 to 34 years of age, whereas the highest proportion of virtual patients were 35 to 49 years, and 50 to 64 years in the PCP group (Table 2). The majority of patients in all groups were women and had a low disease burden, indicated by a DCI score of zero. Of those with comorbidities, the most common was hypertension. While still low, DCI score and common comorbidities were highest among the PCP group (12% of the PCP group had DCI score of 1 or 2, compared with 10% of virtual and 7% of other groups).

Table 2. Baseline characteristics.

	Virtual, 4635	RHC[a], 13,832		UCC[b], 13,757		ED[c], 13,840		PCP[d], 13,881	
	mean (SD) / n (%)	mean (SD) / n (%)	*P* [e]	mean (SD) / n (%)	*P* [e]	mean (SD) / n (%)	*P* [e]	mean (SD) / n (%)	*P* [e]
Age of adults, mean (SD)	40.1 (10.8)	39.3 (12.7)	<.001	37.5 (13.1)	<.001	38.1 (13.5)	<.001	42.7 (13.2)	<.001
Age of children, mean (SD)	8.4 (5.2)	9.8 (4.7)	<.001	9.4 (5.1)	<.001	7.1 (5.4)	<.001	7.1 (5.1)	<.001
Age category, n (%)			<.001		<.001		<.001		<.001
<18	557 (12.0)	1664 (12.0)		1622 (11.8)		1676 (12.1)		1675 (12.1)	
18-34	1414 (30.5)	4814 (34.8)		5510 (40.1)		5355 (38.7)		3543 (25.5)	
35-49	1729 (37.3)	4248 (30.7)		3912 (28.4)		3784 (27.3)		4188 (30.2)	
50-64	935 (20.2)	3106 (22.5)		2713 (19.7)		3025 (21.9)		4475 (32.2)	
Female, n (%)	2837 (61.2)	9143 (66.1)	<.001	5221 (62.1)	.31	8111 (58.6)	.002	8472 (61.0)	.83
Deyo-Charlson Index Score, n (%)			<.001		<.001		<.001		<.001
0	4174 (90.1)	12,856 (92.9)		12,802 (93.1)		12,828 (92.7)		12,205 (87.9)	
1	273 (5.9)	540 (3.9)		476 (3.5)		354 (2.6)		911 (6.6)	
2	188 (4.1)	436 (3.2)		479 (3.5)		658 (4.8)		765 (5.5)	
Comorbidities, n (%)									
Diabetes mellitus	127 (2.7)	308 (2.2)	.05	288 (2.1)	.01	329 (2.4)	.17	560 (4.0)	<.001
Hypertension	390 (8.4)	1081 (7.8)	.19	1147 (8.3)	.87	1620 (11.7)	<.001	1986 (14.3)	<.001
Ischemic heart disease	23 (0.5)	91 (0.7)	.19	93 (0.7)	.18	142 (1.0)	.001	184 (1.3)	<.001
Congestive heart failure	4 (0.1)	7 (0.1)	.39	12 (0.1)	.99	20 (0.1)	.34	14 (0.1)	.78
Chronic obstructive pulmonary disease	23 (0.5)	30 (0.2)	.002	32 (0.2)	.005	66 (0.5)	.87	127 (0.9)	.01
Asthma	190 (4.1)	380 (2.7)	<.001	409 (3.0)	<.001	443 (3.2)	.004	596 (4.3)	.70

[a] RHC: retail health clinic.

[b] UCC: urgent care center.

[c] ED: emergency department.

[d] PCP: primary care physician.

[e] *P* values show level of significance of the differences between each location versus virtual visits.

Care Comparisons

Subsequent outpatient medical care after the initial visit was similar between virtual visits and other treatment settings. The percentage of follow-up visits within 3 weeks of the index visit, which is a potential indicator of misdiagnosis or treatment failure, was similar between the virtual (28.09%), RHC (28.59%; *P*=.54), and PCP groups (28.10%; *P*=.96; Table 3). While the UCC group had slightly fewer follow-up visits (25.62%; *P*<.001), the ED group had more (34.19%; *P*<.001). The virtual group had fewer ED visits within 3 weeks of the index visit (1.32%) compared with the UCC (2.68%; *P*<.001), ED (6.47%; *P*<.001), and PCP groups (1.84%; *P*=.02) but similar to the RHC group (1.61%; *P*=.14). The percentage of hospitalizations within 3 weeks followed a similar pattern, with the percentage for the virtual group (0.15%) lower than the UCC (0.41%; *P*=.01), ED (0.96%; *P*<.001), and PCP groups (0.37%; *P*=.02) and similar to the RHC group (0.28%; *P*=.12).

Table 3. Care patterns.

	Virtual	RHC[a]		UCC[b]		ED[c]		PCP[d]	
	n (%)	n (%)	*P* [e]	n (%)	*P* [e]	n (%)	*P* [e]	n (%)	*P* [e]
All-cause follow-up care within 21 days of index visit, all conditions									
Outpatient evaluation and management visit	1302 (28.1)	3955 (28.6)	.51	3525 (25.6)	.001	4732 (34.2)	<.001	3900 (28.1)	.99
ED visit	61 (1.3)	223 (1.6)	.16	368 (2.7)	<.001	895 (6.5)	<.001	255 (1.8)	.02
Inpatient visit	7 (0.2)	39 (0.3)	.12	57 (0.4)	.01	133 (1.0)	<.001	52 (0.4)	.02
Lab tests within 21 days, all conditions	582 (12.6)	5089 (36.8)	<.001	5367 (39.0)	<.001	7356 (53.2)	<.001	5192 (37.4)	<.001
UTI	85 (20.6)	1085 (87.5)	<.001	1189 (96.0)	<.001	1222 (98.9)	<.001	1095 (88.4)	<.001
Pharyngitis	45 (15.8)	770 (90.2)	<.001	719 (84.3)	<.001	560 (65.8)	<.001	627 (73.5)	<.001
Sinusitis	185 (11.0)	949 (18.8)	<.001	1243 (24.6)	<.001	2351 (46.8)	<.001	1302 (25.7)	<.001
Bronchitis	40 (10.1)	285 (23.9)	<.001	271 (22.7)	<.001	648 (54.6)	<.001	308 (25.8)	<.001
Imaging rates within 21 days, all conditions	307 (6.6)	826 (6.0)	.11	1207 (8.8)	<.001	5960 (43.1)	<.001	1563 (11.3)	<.001
Cough	18 (11.4)	46 (9.7)	.55	106 (22.4)	.003	397 (84.6)	<.001	111 (23.5)	.001
Bronchitis	34 (8.6)	114 (9.6)	.59	193 (16.2)	<.001	844 (71.0)	<.001	212 (17.8)	<.001
UTI	34 (8.2)	85 (6.9)	.35	132 (10.7)	.16	763 (61.7)	<.001	227 (18.3)	<.001
URI	69 (8.1)	144 (5.7)	.01	203 (8.0)	.91	1067 (42.1)	<.001	236 (9.3)	.31
Sinusitis	90 (5.3)	287 (5.7)	.59	358 (7.1)	.01	2152 (42.8)	<.001	497 (9.8)	<.001
Antibiotic fill rates within 3 days[f]									
Any of the 6 infections below	1918 (70.5)	4193 (64.2)	<.001	4243 (67.9)	.02	3534 (56.7)	<.001	4477 (68.2)	.03
Sinusitis	971 (83.9)	2340 (86.3)	.06	2084 (79.2)	.001	1835 (67.8)	<.001	2327 (82.9)	.42
Pharyngitis	130 (74.3)	138 (29.6)	<.001	236 (53.8)	<.001	199 (46.4)	<.001	249 (53.7)	<.001
Bronchitis	191 (68.5)	278 (40.8)	<.001	521 (76.4)	.01	393 (62.1)	.06	545 (78.1)	.002
Conjunctivitis	157 (63.8)	463 (78.6)	<.001	363 (64.9)	.76	278 (51.8)	.002	373 (61.1)	.47
UTI	217 (76.4)	628 (90.5)	<.001	473 (74.0)	.44	419 (65.6)	.001	415 (62.6)	<.001
URI	252 (43.5)	346 (24.9)	<.001	566 (43.7)	.30	410 (31.9)	<.001	568 (43.0)	.82
Antibiotic type									
Broad-spectrum antibiotic as first-line treatment[g]	1219 (69.0)	2299 (60.0)	<.001	2561 (66.3)	.04	1961 (61.9)	<.001	2704 (69.3)	.82

[a] ED: emergency department.

[b] PCP: primary care physician.

[c] RHC: retail health clinic.

[d] UCC: urgent care center.

[e] *P* values show level of significance of the differences between each location versus virtual visits.

[f] Sample includes patients with the condition of interest and pharmacy coverage.

[g] Sample includes patients with antibiotics fill without history of antibiotic use in prior 60 days.

Overall laboratory tests within 3 weeks of the index date (including during the initial visit for nonvirtual visits) were lower for the virtual group (12.56%) compared with RHC (36.79%; *P*<.001), UCC (39.01 %; *P*<.001), ED (53.15%; *P*<.001), and PCP (37.40%; *P*<.001; Table 3). Lab testing was particularly low in virtual episodes compared with other locations of care for pharyngitis and UTI, where testing for a bacterial infection is common. Overall imaging rates were similar between the virtual and RHC groups (6.62% vs 5.97%, respectively; *P*=.11), but much lower than ED (43.06%; *P*<.001), and somewhat lower than UCC (8.77%; *P*<.001) and PCP groups (11.26%; *P*<.001). While lab and imaging rates differed by condition, the pattern of lower rates of lab and imaging testing in virtual episodes (except similar rates to RHC imaging) was consistent across conditions (see Table 3).

Overall antibiotic fills within 3 days for the 6 most commonly treated infections (excluding influenza) was somewhat higher in the virtual group (70.51%) compared with all other sites (RHC 64.18%; *P*<.001, UCC 67.94%; *P*=.02, ED 56.73%; *P*<.001, PCP 68.19%; *P*=.03; Table 3), although there was variation by infection type. Fill rates after virtual visits tended to follow more similar patterns to UCC (for conjunctivitis, URI, and UTI) and PCP (for conjunctivitis, sinusitis, and URI) than RHC (similarity to sinusitis only) and ED (similarity to bronchitis only). Antibiotic fills were substantially higher after virtual visits than all other locations for pharyngitis (virtual 74.3%, RHC 29.6%; *P*<.001, UCC 53.8%; *P*<.001, ED 46.4%; *P*<.001, PCP 53.7%; *P*<.001).

Broad-spectrum antibiotics were used as first-line treatment in the virtual group (68.99%) at a similar rate to the PCP group (69.28%; *P*=.82), but more often than in the RHC (59.98%; *P*<.001), UCC (66.28%; *P*=.04), and ED groups (61.90%; *P*<.001).

Cost Comparisons

Total costs per episode were $36, $153, $1735, and $162 more expensive at RHC, UCC, ED, and PCP settings, respectively, compared with virtual visits (Tables 4 and 5).

Table 4. Cost of retail health clinic and urgent care center visits compared with virtual visits, adjusted for age categories and baseline comorbidities.

	Virtual		RHC			UCC		
	n	Mean, $	n	Mean, $ (95% CI)[a]	Relative (95% CI)[b]	n	Mean, $ (95% CI)[a]	Relative (95% CI)[b]
Index visit	4635	49	13,832	74 (72-75)	1.52 (1.49-1.54)	13,757	134 (131-136)	2.75 (2.70-2.79)
Follow-up, medical	4635	200	13,832	204 (189-220)	1.02 (0.95-1.10)	13,757	266 (247-287)	1.33 (1.23-1.43)
Pharmacy	3182	90	7518	97 (91-104)	1.08 (1.01-1.15)	7188	92 (86-98)	1.03 (0.96-1.09)
Total (sum, estimate)		339		375	1.11		492	1.45

[a] Mean cost, adjusted to virtual visit distribution of age and comorbidities.

[b] Relative = ratio of how much more expensive RHC/UCC visits are compared with virtual visits after adjustments.

Table 5. Cost of emergency department and primary care physician visits compared with virtual visits, adjusted for age categories and baseline comorbidities.

	Virtual		ED			PCP		
	n	Mean, $	n	Mean, $ (95% CI)[a]	Relative (95% CI)[b]	n	Mean,$ (95% CI)[a]	Relative (95% CI)[b]
Index visit	4635	49	13,840	1404 (1381-1428)	28.87 (28.39-29.36)	13,881	109 (107-111)	2.25 (2.21-2.28)
Follow-up, medical	4635	200	13,840	584 (542-631)	2.92 (2.70-3.15)	13,881	288 (267-311)	1.44 (1.33-1.55)
Pharmacy	3182	90	7227	86 (81-92)	0.96 (0.90-1.02)	7629	104 (97-110)	1.15 (1.08-1.23)
Total (sum, estimate)		339		2074	6.12		501	1.48

[a] Mean cost, adjusted to virtual visit distribution of age and comorbidities.

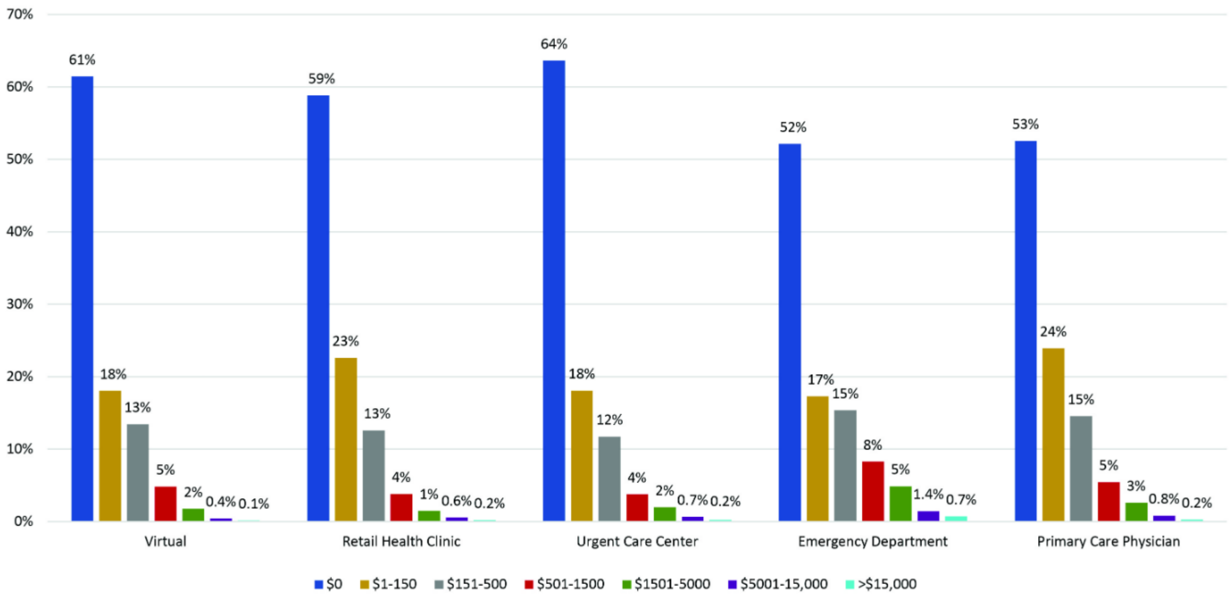
[b] Relative = ratio of how much more expensive ED/PCP visits are compared with virtual visits after adjustments.

As expected, the adjusted mean cost of the initial visit was lower for the virtual group ($49) than for RHC ($74; *P*<.001), UCC ($134; *P*<.001), ED ($1404; *P*<.001) and PCP ($109; *P*<.001; Tables 4 and 5). Follow-up medical costs for the virtual group were similar to or lower than the costs for each of the other sites of care. In the virtual group, 61.47% of patients had no follow-up medical costs, compared with 58.82% RHC, 63.63% UCC, 52.14% ED, and 52.53% PCP (Figure 2). Follow-up costs exceeded $500 for 7.08% of patients in the virtual group, compared with 6.01% RHC, 6.59% UCC, 15.23% ED, and 9.04% PCP; follow-up costs exceeded $5000 for 0.54% in the virtual group compared with 0.74% RHC, 0.88% UCC, 2.11% ED, and 1.03% PCP. Adjusted average follow-up medical costs were similar between the virtual ($200) and RHC groups ($204; *P*=.62) but higher for the UCC ($266; *P*<.001), ED ($584; *P*<.001), and PCP groups ($288; *P*<.001). At $90, the adjusted average pharmacy cost for a virtual episode was similar to UCC

($92; *P*=.44) and ED ($86; *P*=.21) and somewhat lower than RHC ($97; *P*=.02) and PCP ($104; *P*<.001).

While average episode costs differed by condition, they tended to follow a similar pattern of virtual visits having lower medical costs than care at other locations across conditions (Multimedia Appendix 2). Additionally, the sensitivity analysis with winsorized values showed consistent relative costs of care between virtual and the other locations as the original analysis, except for lower RHC and UCC follow-up medical costs, although total episode cost differences were consistent (Multimedia Appendix 3).

Figure 2. Follow-up medical costs, unadjusted.



Discussion

Principal Findings

This retrospective, real-world analysis demonstrated that care received through virtual visits for nonurgent conditions was comparable to that received in in-person health care settings. Patients receiving care through virtual visits had similar follow-up outpatient evaluation and management visit rates as patients using other locations. This finding suggests not only that patients using virtual visits had their health problems resolved at similar rates as patients treated at other locations but also that patients were not using virtual visits as a first step before seeking in-person care. Interestingly, follow-up visit rates for the virtual group mirrored patients' self-reported resolution of symptoms. An informal survey administered as part of the health plan's virtual care program found 79% of patients who used it reported complete resolution of their health care concerns (personal communication, W Adamson).

Lab testing rates, both overall and at the individual diagnosis level, were lower during virtual visits episodes than all in-person settings. Lab testing may be high at in-person locations for some conditions where it may not be needed to confirm the patient's diagnosis [17,18]. However, the large differences between lab testing rates for pharyngitis and UTI suggest that patients with virtual care visits may not receive testing for these conditions where differentiating between a viral and bacterial infection is important for treatment. Antibiotics for the infections most commonly diagnosed through virtual visits were prescribed significantly more often after virtual visits than any of the other in-person treatment settings. The difference between virtual visits and other settings was even greater for pharyngitis, presumably because streptococcal infection was not ruled out with a lab test. While the use of broad-spectrum antibiotics was similar between virtual visits and PCP, it was higher than the other in-person settings.

Episodes for patients who sought care at any of the in-person settings were more expensive than similar episodes beginning with a virtual visit. In addition to the virtual visit itself being less expensive than in-person visits, follow-up medical costs were lower after virtual visits than all other locations except for RHCs. Some of the lower episode costs can be attributed to lower rates of ED or inpatient follow-up care in addition to lower laboratory and imaging rates during the episode.

Limitations

A unique strength of this study was the large database allowing for a 3:1 match of episode-based care received in a number of alternative settings, but the exclusive use of claims data introduced several limitations. The accuracy of the diagnosis in claims may be a particular concern for virtual visits, where it may be more difficult for providers to diagnose a condition without a physical examination or supporting laboratory tests. Such errors not only may lead to inaccuracies in cost comparisons, but may also affect care patterns. It is not possible to determine disease severity from a diagnosis code alone, so cases seen in the ED, for example, may have been more severe, requiring more treatment than an average case handled by a virtual visit. Furthermore, claims do not provide complete information on the reasons patients chose a specific site of care.

While patients may have chosen the ED because they perceived it to be the most convenient option even for a minor illness, it is also possible they believed their condition was severe and required urgent medical attention. However, the conditions included in this study tend to be relatively minor and treatable in nonurgent settings. Additionally, patients who chose virtual visits may have differed from those who chose other treatment settings in terms of their perception of the urgency of their condition, their health literacy, or their level of comfort using computers [7,8]. The analysis did not take into account whether a patient had multiple diagnoses at a single visit, which may have led to a more costly visit or additional follow-up care. Follow-up visits and costs may or may not have been related to the original visit, and we were unable to determine conclusively if a complaint was resolved at the original visit, if the follow-up visit was part of appropriate care, or if the follow-up visit represented inappropriate care-seeking behavior on the part of the patient. Additionally, claims do not provide sufficient information to determine if virtual visits were used in situations when individuals would have otherwise waited to see if the problem resolved on its own. A recent study suggested that RHCs increased utilization for low-acuity visits due to their lower price or convenience [19]. However, whether this is good or bad is a matter of perspective, since while costs may increase, the alternative care options may allow patients who may not have otherwise received appropriate care to receive the care they needed. Hence, our study focused on the difference in cost between care options, rather than potential savings of introducing virtual visits into the market or the necessity of the visits.

Comparison With Prior Studies

The rate of antibiotic prescriptions for the conditions included here may warrant additional study. Based on current guidelines and Choosing Wisely recommendations [20-22], prescription rates may be higher than desired in a variety of care settings. The virtual visits were associated with somewhat higher rates of antibiotic prescriptions than other sites of care overall, including for conditions for which clinical guidelines typically do not recommend antibiotics as a first line of treatment [20-22], although results were mixed when considered by condition. In some cases, the higher antibiotic prescription rate may have been due to a lower rate of laboratory testing associated with virtual visits, since a bacterial versus viral diagnosis could not be confirmed by lab test during virtual visits (eg, pharyngitis, where lab test rates were particularly low for virtual visits compared to other locations). This finding is consistent with previous research, which found higher antibiotic prescribing rates for telehealth than for office visits, especially for pharyngitis, bronchitis, and UTIs [23-25]. Future telehealth programs, particularly those integrated with a medical home or used for patients with an already existing physician-patient relationship, may be able to develop workflows that incorporate lab testing and may help with antibiotic prescribing decisions.

Conclusions

Virtual visits are growing rapidly, and our results indicate they are inexpensive alternatives to acute care administered at other locations. Patients receiving care through virtual visits seemed to have adequate clinical resolution compared with patients receiving care elsewhere, based on follow-up visit rates. Patients did receive additional care, such as laboratory testing or imaging, presumably when needed. Virtual visits did not appear to add to the total amount of care received as part of a care episode, as patients did not often seek care through telehealth plus another site for the same condition.

Expansion of virtual health care services is inevitable given the growing use of mobile devices, patient demand for immediate and convenient access to care, and the continuously growing demands on physicians' time. The focus of further research on virtual health care should be about optimizing patient outcomes for conditions best suited for virtual visits and examining how virtual visits can be used by physicians who have an existing personal relationship with the patient.

Acknowledgments

The authors thank D Marc Cram and Dianna Hayden for programming support and Cheryl Jones for editorial assistance. DM Cram, D Hayden, and C Jones are employees of HealthCore, Inc.

Funding for this study was provided by Anthem, Inc. The authors are solely responsible for the study design, data collection, and interpretation; approvals from Anthem were not required for submission.

Conflicts of Interest

A Gordon and A DeVries are employees of HealthCore, Inc, a wholly owned outcomes research subsidiary of Anthem, Inc. W Adamson is an employee of Anthem, Inc, who works on LiveHealth Online. The authors were not compensated for this study beyond their salaries.

Multimedia Appendix 1

International Classification of Diseases, Ninth Revision, diagnosis codes.

[XLSX File (Microsoft Excel File), 41KB - jmir_v19i2e35_app1.xlsx]

Multimedia Appendix 2

Adjusted mean index visit medical, follow-up medical, and pharmacy costs, by condition.

[XLSX File (Microsoft Excel File), 53KB - jmir_v19i2e35_app2.xlsx]

Multimedia Appendix 3

Comparing original and winsorized adjusted costs (sensitivity analysis).

[XLSX File (Microsoft Excel File), 9KB - jmir_v19i2e35_app3.xlsx]

References

1. Mehrotra A. The convenience revolution for treatment of low-acuity conditions. JAMA 2013 Jul 3;310(1):35-36 [FREE Full text] [doi: 10.1001/jama.2013.6825] [Medline: 23821082]
2. Ashwood JS, Reid RO, Setodji CM, Weber E, Gaynor M, Mehrotra A. Trends in retail clinic use among the commercially insured. Am J Manag Care 2011 Nov;17(11):e443-e448 [FREE Full text] [Medline: 22200061]
3. DeVries A, Li C, Oza M. Strategies to reduce nonurgent emergency department use: experience of a Northern Virginia Employer Group. Med Care 2013 Mar;51(3):224-230. [doi: 10.1097/MLR.0b013e3182726b83] [Medline: 23222497]
4. Hersh WR, Wallace JA, Patterson PK, Shapiro SE, Kraemer DF, Eilers GM, et al. Telemedicine for the Medicare population: pediatric, obstetric, and clinician-indirect home interventions. Evid Rep Technol Assess (Summ) 2001 Aug(24 Suppl):1-32 [FREE Full text] [Medline: 11569328]
5. 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]
6. Daniel H, Sulmasy LS, Health and Public Policy Committee of the American College of Physicians. Policy recommendations to guide the use of telemedicine in primary care settings: an American College of Physicians position paper. Ann Intern Med 2015 Nov 17;163(10):787-789. [doi: 10.7326/M15-0498] [Medline: 26344925]
7. Albert SM, Shevchik GJ, Paone S, Martich GD. Internet-based medical visit and diagnosis for common medical problems: experience of first user cohort. Telemed J E Health 2011 May;17(4):304-308 [FREE Full text] [doi: 10.1089/tmj.2010.0156] [Medline: 21457013]
8. Hickson R, Talbert J, Thornbury WC, Perin NR, Goodin AJ. Online medical care: the current state of “eVisits” in acute primary care delivery. Telemed J E Health 2015 Feb;21(2):90-96. [doi: 10.1089/tmj.2014.0022] [Medline: 25474083]
9. Jung C, Padman R. Virtualized healthcare delivery: understanding users and their usage patterns of online medical consultations. Int J Med Inform 2014 Dec;83(12):901-914. [doi: 10.1016/j.ijmedinf.2014.08.004] [Medline: 25193501]
10. Mehrotra A, Paone S, Martich GD, Albert SM, Shevchik GJ. Characteristics of patients who seek care via eVisits instead of office visits. Telemed J E Health 2013 Jul;19(7):515-519 [FREE Full text] [doi: 10.1089/tmj.2012.0221] [Medline: 23682589]
11. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Sep 07(9):CD002098. [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]
12. Mehrotra A, Paone S, Martich GD, Albert SM, Shevchik GJ. A comparison of care at e-visits and physician office visits for sinusitis and urinary tract infection. JAMA Intern Med 2013 Jan 14;173(1):72-74 [FREE Full text] [doi: 10.1001/2013.jamainternmed.305] [Medline: 23403816]
13. Uscher-Pines L, Mulcahy A, Cowling D, Hunter G, Burns R, Mehrotra A. Access and quality of care in direct-to-consumer telemedicine. Telemed J E Health 2016 Apr;22(4):282-287. [doi: 10.1089/tmj.2015.0079] [Medline: 26488151]
14. Uscher-Pines L, Mulcahy A, Cowling D, Hunter G, Burns R, Mehrotra A. Antibiotic prescribing for acute respiratory infections in direct-to-consumer telemedicine visits. JAMA Intern Med 2015 Jul;175(7):1234-1235. [doi: 10.1001/jamainternmed.2015.2024] [Medline: 26011763]
15. LiveHealth Online. 2015. URL: https://www.livehealthonline.com/ [accessed 2017-01-26] [WebCite Cache ID 6kgvymb2k]
16. Mehrotra A, Liu H, Adams JL, Wang MC, Lave JR, Thygeson NM, et al. Comparing costs and quality of care at retail clinics with that of other medical settings for 3 common illnesses. Ann Intern Med 2009 Sep 1;151(5):321-328 [FREE Full text] [Medline: 19721020]
17. Albert RH. Diagnosis and treatment of acute bronchitis. Am Fam Physician 2010 Dec 01;82(11):1345-1350 [FREE Full text] [Medline: 21121518]
18. Rosenfeld RM, Piccirillo JF, Chandrasekhar SS, Brook I, Kumar KA, Kramper M, et al. Clinical practice guideline (update): Adult Sinusitis Executive Summary. Otolaryngol Head Neck Surg 2015 Apr;152(4):598-609. [doi: 10.1177/0194599815574247] [Medline: 25833927]
19. Ashwood JS, Gaynor M, Setodji CM, Reid RO, Weber E, Mehrotra A. Retail clinic visits for low-acuity conditions increase utilization and spending. Health Aff (Millwood) 2016 Mar;35(3):449-455. [doi: 10.1377/hlthaff.2015.0995] [Medline: 26953299]

20. American Academy of Family Physicians. Don't prescribe antibiotics for otitis media in children aged 2-12 years with non-severe symptoms where the observation option is reasonable. 2016. Choosing Wisely URL: http://www.choosingwisely.org/clinician-lists/american-academy-family-physicians-antibiotics-for-otitis-media-in-children/ [WebCite Cache ID 6kgwliDnA]
21. American Academy of Pediatrics. Antibiotics should not be used for apparent viral respiratory illnesses (sinusitis, pharyngitis, bronchitis). 2016. Choosing Wisely URL: http://www.choosingwisely.org/clinician-lists/american-academy-pediatrics-antibiotics-for-children-with-viral-respiratory-illness [WebCite Cache ID 6kgx1ojHm]
22. Infectious Diseases Society of America. Avoid prescribing antibiotics for upper respiratory infections. 2016. Choosing Wisely URL: http://www.choosingwisely.org/clinician-lists/infectious-diseases-society-antbiotics-for-upper-respiratory-infections/ [accessed 2016-09-21] [WebCite Cache ID 6kgxYCGPo]
23. Harris AM, Hicks LA, Qaseem A, High Value Care Task Force of the American College of Physicians and the Centers for Disease Control and Prevention. Appropriate antibiotic use for acute respiratory tract infection in adults: advice for high-value care from the American College of Physicians and the Centers for Disease Control and Prevention. Ann Intern Med 2016 Mar 15;164(6):425-434. [doi: 10.7326/M15-1840] [Medline: 26785402]
24. Wong DM, Blumberg DA, Lowe LG. Guidelines for the use of antibiotics in acute upper respiratory tract infections. Am Fam Physician 2006 Sep 15;74(6):956-966 [FREE Full text] [Medline: 17002029]
25. American Academy of Family Physicians. 2016. Recommendations by type: clinical practice guidelines URL: http://www.aafp.org/patient-care/clinical-recommendations/clinical-practice-guidelines.html [WebCite Cache ID 6kh1AzsRL]

Abbreviations

CPT: Current Procedural Terminology
DCI: Deyo-Charlson Comorbidity Index
ED: emergency department
HIRD: HealthCore Integrated Research Database
ICD-9: International Classification of Diseases, Ninth Revision
NPI: National Provider Identifier
PCP: primary care physician
RHC: retail health clinic
UCC: urgent care center
URI: upper respiratory infection
UTI: urinary tract infection

Edited by G Eysenbach; submitted 11.10.16; peer-reviewed by AJ Goodin, S Kurklinsky; comments to author 24.11.16; revised version received 03.01.17; accepted 21.01.17; published 17.02.17.

Please cite as:
Gordon AS, Adamson WC, DeVries AR
Virtual Visits for Acute, Nonurgent Care: A Claims Analysis of Episode-Level Utilization
J Med Internet Res 2017;19(2):e35
URL: http://www.jmir.org/2017/2/e35/
doi:10.2196/jmir.6783
PMID:28213342

Original Paper

Public Funding and Open Access to Research: A Review of Canadian Multiple Sclerosis Research

Caitlin Bakker[1], MLIS; Carol Stephenson[2], MLIS; Erin Stephenson[3], PhD; Debbie Chaves[4], MLIS, PhD

[1]Health Sciences Libraries, University of Minnesota, Minneapolis, MN, United States

[2]Council of Prairie and Pacific University Libraries, Vancouver, BC, Canada

[3]Hotchkiss Brain Institute and the Department of Clinical Neurosciences, University of Calgary, Calgary, AB, Canada

[4]Library, Wilfrid Laurier University, Waterloo, ON, Canada

Corresponding Author:
Caitlin Bakker, MLIS
Health Sciences Libraries
University of Minnesota
303 Diehl Hall, 505 Essex Street SE
Minneapolis, MN,
United States
Phone: 1 612 301 1353
Fax: 1 612 626 2454
Email: cjbakker@umn.edu

Abstract

Background: Multiple sclerosis (MS), a progressive demyelinating disease of the brain and spinal cord, is the leading cause of nontraumatic neurological damage in young adults. Canada has one of the highest reported incidents of MS, with estimates between 55 and 240 per 100,000 individuals. Between 2009 and 2014, the MS Society of Canada provided over Can $90 million to researchers and, since 2013, has encouraged researchers to make both current and previous research products openly available.

Objective: The goal of the study was to determine the open access (OA) cost implications and repository policies of journals frequently used by a sample of MS researchers. This study benchmarked current publishing preferences by MS Society of Canada researchers by examining the OA full-text availability of journal articles written by researchers funded between 2009 and 2014.

Methods: Researchers were identified from the 2009 to 2014 annual MS Society of Canada Research Summaries. Articles were identified through searches in Web of Science, Scopus, Medline and Embase (both via OVID). Journal level analysis included comparison of OA policies, including article processing charges (APCs) and repository policies. Data were analyzed using descriptive statistics.

Results: There were 758 articles analyzed in this study, of which 288 (38.0%) were OA articles. The majority of authors were still relying on journal policies for deposit in PubMed Central or availability on publisher websites for OA. Gold OA journals accounted for 10.2% of the journals in this study and were associated with significantly lower APCs (US $1900) than in hybrid journals (US $3000). Review of the journal self-archiving options highlighted the complexity of stipulations that authors would have to navigate to legally deposit a version of their article.

Conclusions: This study found that there are currently researcher- and publisher-imposed barriers to both the gold and green roads to OA. These results provide a current benchmark against which efforts to enhance openness can be measured and can serve as a reference point in future assessments of the impact of OA policies within this field. With funding agencies worldwide releasing OA mandates, future success in compliance will require changes to how researchers and publishers approach production and dissemination of research.

(J Med Internet Res 2017;19(2):e52) doi:10.2196/jmir.6250

KEYWORDS

multiple sclerosis; open access publishing; research support as topic

Introduction

Multiple sclerosis (MS), a progressive demyelinating disease of the brain and spinal cord, is the leading cause of nontraumatic neurological damage in young adults [1]. Depending on the affected areas of the brain, balance, vision, hearing, thinking, and memory may ultimately be impacted. Canada has one of the highest reported incidents of MS, with estimates between 55 and 240 per 100,000 individuals [2]. Despite over 200 years of research and significant recent findings [3], a standard, evidence-based treatment proven to halt the chronic progression and long-term disability has remained elusive. Given the debilitating character of MS and its prevalence in Canada, it is not surprising that the MS Society of Canada infuse significant funding to advance research. Between 2009 and 2014, the MS Society of Canada provided over Can $90 million to researchers in four major research areas: (1) symptom management and quality of life, (2) progression and therapies, (3) cause and risk factors, and (4) nerve damage and repair [4].

The MS Society of Canada has strongly encouraged open access (OA) and the broad dissemination of research. The MS Society of Canada's OA policy, which came into effect in July 2013, requires grant recipients to make every effort to ensure their peer-reviewed publications are available OA within 6 months of publication [5]. While the policy came into effect in 2013, the MS Society of Canada encourages retroactive compliance for research funded before that date.

The information needs and information-seeking behaviors of clinicians and patients with MS have been well documented [6-10]. Studies have shown that MS patients are demanding active roles in their treatments and the decision-making process [11-13], and health care information on the Internet ranks second to health professionals as a source of information for patients [14]. Although scientists and funding agencies recognize the potential public value of research and of making work available OA, resources are required to facilitate the process. Whether these are financial resources to pay for article processing charges (APCs) or the time to negotiate with journals to allow for deposit in an appropriate repository, OA requires work on the part of the researcher. The appropriate allocation of resources for broad dissemination remains a question.

Using the MS Society of Canada as a case study, this paper examines the OA full-text availability of scientific articles by Canadian researchers funded between 2009 and 2014, including the publishing venues and associated costs of making work openly available. The goal of the study was to determine the extent of OA, the cost implications, and repository policies of journals frequently used by MS researchers as an assessment of the current context which can be used in future evaluations of the extent of openness and the effectiveness of mandates in making research OA.

Methods

Database Search

Researchers were identified from the 2009 to 2014 annual MS Society of Canada Research Summaries [15]. Researcher names and affiliations, titles of the research projects, funded amounts, and funded years were recorded in an Excel (Microsoft) file. The name of the researcher and the keyword "Multiple Sclerosis" were searched in 4 databases: Web of Science, Scopus, and Embase and Medline (both via OVID platform). Due to the inconsistencies in funding information in journal metadata, searches did not incorporate the MS Society of Canada as a funding body. In selected cases, an author affiliation was used to aid in disambiguation. Publications were limited to the first year the researcher received MS Society funding and all subsequent years. Articles published in 2015 were included in this analysis to account for the time necessary to finalize research, write publications, and complete the publishing process. Publication types were limited to journal articles, review articles, and conference papers reproduced in their entirety in a journal, including items that were both published and in press. Searches were limited to English language only.

The results from the database searches were merged into a single, deduplicated file of articles. A secondary quality check was performed by manually reviewing the article details to ensure relevance and accuracy. The article title was searched in Google and PubMed to determine if the full-text was openly accessible (open access) either through the publisher website or PubMed Central. The journal policy for access through PubMed Central and the version of the article on PubMed Central (author manuscript to meet compliance with a funder policy or a journal OA policy) were recorded. Full-text access through other sources such as Academia.edu or ResearchGate was not included because these sites do not necessarily guarantee that the full-text option is a legitimate copy.

Colors to Categorize Policies

Subscription policies were used to identify the title as a pure gold journal or a hybrid journal. The self-archiving policies for each journal were obtained from SHERPA RoMEO, a UK academic supported database that provides information regarding copyright policies and rights retained by the authors when publishing in specific academic journals. SHERPA RoMEO uses colors to categorize publishers' archiving policies: white (archiving is not formally supported), yellow (authors can archive preprint, that is, prerefereeing), blue (authors can archive postprint, that is, final draft postrefereeing, or publisher's version), or green (authors can archive preprint and postprint, or publisher's version). Each journal in the study was assigned one of the 4 colors. The OA policies were analyzed first by publisher, per journal, and then at the article level.

For each journal, APC policies were searched on the publisher website. APC costs were also converted to US dollars. The maximum APC fee quoted by the journal was recorded where there were variable fees listed. APCs fees were variable depending on the type of article, memberships, institutional affiliations, funding bodies, or which Creative Commons license was selected.

Results

Overview of Publishing Venues

The MS Society of Canada funded 77 Canadian researchers between 2009 and 2014 in one of their four major research areas. This study identified a total of 758 MS related articles in 211 journals produced by 61 publishers. Of the 758 articles, 288 (38.0%) were OA articles produced in 93 journals by 42 publishers.

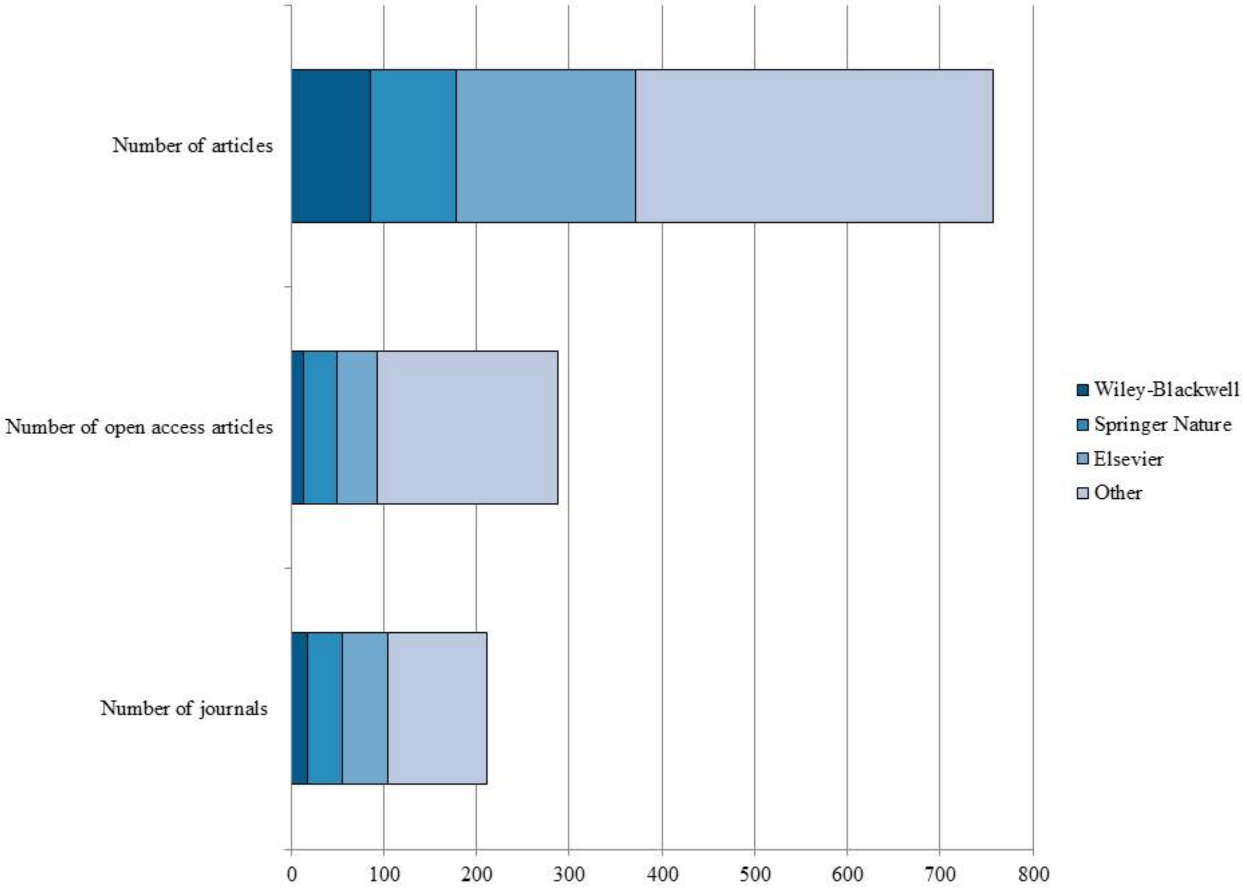
When article distribution per journal was analyzed to identify the preferred or most popular publishing venues, there was a noticeable difference in the dominance of publishers. Three major publishers (Elsevier, Wiley, and Springer Nature) dominated the research output, with 105 of 211 journals (49.7%) and 374 of 758 articles (49.3%). However, these publishers only accounted for 93 of the 288 OA articles (32.2%) in the study (Figure 1). Only one Elsevier journal (*NeuroImage: Clinical*) and 3 Springer journals (*BMC Health Services Research*, *BMC Neurology*, and *Journal of Neuroinflammation*) were pure gold in which every article was OA full-text upon publication. There were no pure gold journals from Wiley-Blackwell identified in the articles analyzed in this study.

A total of 19 journals from 13 publishers accounted for 388 of the 758 articles (51.2%) and 146 of the 288 OA articles (50.7%) in the study (Table 1). The three major publishers (Elsevier, Wiley, and Springer Nature) published 8 of the 19 journals in this preferred grouping, but SAGE and Lippincott Williams and Wilkins (LWW) were the top publishers with both their journals, *Multiple Sclerosis* (SAGE) and *Neurology* (LWW) accounting for 142 of the 758 articles (18.7%) in the study. Only one pure gold journal (*PLoS ONE*) was in this grouping, ranking 11th, with 14 of the 758 articles in the study (Table 1).

Table 1. The most frequently used journals, accounting for 51.2% of the articles. Number of articles, number of open access (OA) articles, article processing charges (APC), and SHERPA RoMEO color are provided.

Journal title	Articles (n)	OA articles (n)	Maximum APC (US $)	SHERPA RoMEO color
Multiple Sclerosis	76	15	3000	Green
Neurology	66	35	3100	Yellow
Journal of the Neurological Sciences	23	0	2500	Green
Annals of Neurology	23	4	3000	Yellow
GLIA	19	2	3000	Yellow
Journal of Neuroimmunology	18	2	2500	Green
Multiple Sclerosis and Related Disorders	18	2	2500	Green
Journal of Neurology	17	7	3000	Green
Journal of Immunology	16	14	3000	Blue
Lancet Neurology	16	0	5000	Green
PLoS ONE	14	14	1350	Green
Canadian Journal of Neurological Sciences	13	0	2000	Green
NeuroImage	13	2	2200	Green
Journal of Neurology, Neurosurgery and Psychiatry	10	7	2822	Green
Journal of Neuroscience	10	10	2820	Yellow
International Journal of MS Care	9	9	n/a	Ungraded
Neuro-epidemiology	9	6	3250	Green
Brain	9	7	3200	Yellow
JAMA Neurology (Archives of Neurology)	9	8	n/a	White

Figure 1. The top three publishers by journal and article count.



Article Processing Charges

A closer examination of the reason for the availability of OA revealed limited use of APC payments or deposit of author manuscripts in PubMed Central, even though most journals offered these as options (see Figure 2). Of the 288 OA articles, 180 (62.5%) were made available in their final typeset form, either through deposit in PubMed Central (163/180, 90.6%) or through the publisher's website (17/180, 9.4%) without evidence of APC payment. In comparison, author directed OA through APCs accounted for 77 of the 288 articles (26.7%), whereas 31 of the 288 articles (10.8%) were made available as manuscripts deposited in PubMed Central. The limited use of APCs is even more striking when all articles in the study are considered. Whereas 662 of the 758 articles were published in journals with APC options, only 77 of the 758 (10.2%) have been published using that option.

APCs for pure gold and hybrid journals revealed significant differences in potential costs to authors. Of the 211 journals in this study, 22 were pure gold, with each of the 60 articles produced by these journals freely available full-text upon publication. Pure gold journals not only use APCs to cover publication costs but also use institutional memberships or other funding options to provide all content openly accessible upon publications. The pure gold journals in this study had variable APCs depending on memberships, government support, and the type of article published. APCs were listed on publisher sites for 19 of the 22 journals. One of these unlisted journals (*Preventing Chronic Disease*) was produced by the Centers for Disease Control and Prevention which does not use APCs. *Functional Neurology* and *European Neurological Review* also did not list APCs in their author instructions. For the 19 journals for which APCs were found, the average and mode fees were US $1900.

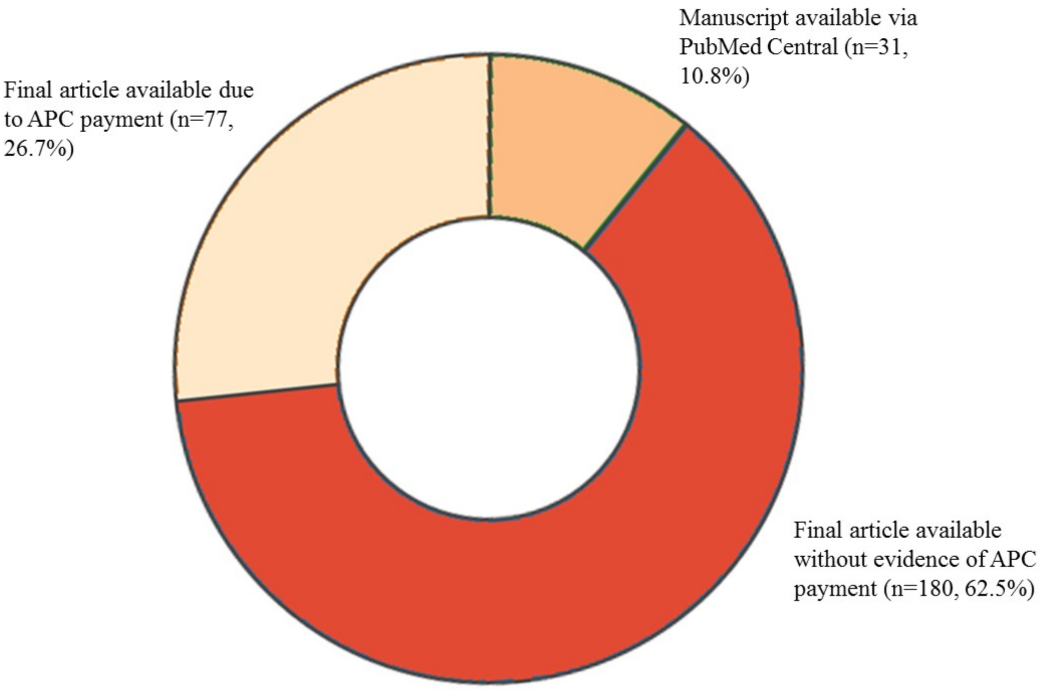
APCs were found for 137 of the 189 hybrid journals. The other 52 journals either required authors to contact the publisher for fee information or it was not evident if the option existed. Fees within a journal could vary depending on memberships, institutional affiliations, funding bodies, or which Creative Commons license was selected. The APCs ranged from US $600 to US $5000. The APC fees were notably higher for hybrid journals than for pure gold journals. The average cost was US $2800, and the mode was US $3000. The APCs for the hybrid journals with the highest number of articles in this study ranged from US $1000 for members in the *Canadian Journal of Neurological Sciences* to US $5000 for *Lancet Neurology*. Table 2 provides a profile of the potential APC costs for articles from this study published in 2014.

Table 2. Snapshot of maximum article processing charges (APCs) for articles published in 2014.

Journal Type	N		APC[a] Information		
	Journals	Articles	Range	Average	Mode
Journals without APC information	18	22	N/A	N/A	N/A
Gold journals with APCs	9	16	US $1350 to US $2310	US $1735	US $1500
Hybrid journals with APCs	54	108	US $600 to US $5000	US $2979	US $3000

[a] APC: article processing charge.

Figure 2. The number of final articles available as open access due to article processing charge (APC) payments, article manuscript deposit in PubMed Central, and final articles available without evidence of APC payment through deposit of published articles to PubMed Central or availability from the publisher website.



Policies for and Prevalence of Self-Archiving

With the impracticability of the APC model for many researchers, the policies for self-deposit were reviewed for the 211 journals in the study. The SHERPA RoMEO website provided information on journal policies for 203 of 211 journals in the study. The ability for authors to self-archive a preprint or postprint of the article to an institutional or subject-based repository was offered by the majority of journals, but there was large variability in the version of the manuscript that could be deposited.

Of the 211 journals, 120 allowed deposit of the final draft (postprint) of the manuscript, 71 allowed the final draft but with additional restrictions, and 11 did not allow the final draft and information was not provided the other 9 journals. The SHERPA RoMEO analysis identified deposit embargoes on 138 of the journals that would have affected 564 of the 758 articles in this study (74%). In addition, the restrictive self-archiving policies (yellow and white SHERPA RoMEO categories) were applied to journals produced by 16 societies.

Of the 77 MS Society of Canada funded researchers included in this study, 62 had affiliations that would offer access to an institutional repository (IR). A secondary search of repositories for the 760 articles found that 8 of the 760 articles (1%) identified in this study were made available through an IR. Of these, all 8 were originally published in pure gold journals, namely in BioMed Central titles. IRs did not appear to function as a mechanism to make otherwise inaccessible content publicly available.

Discussion

Principal Findings

Despite encouragement from the public and the clinical community [16-18], and the MS Society of Canada's goal of making both current and previous research publicly available, the overall rates of openness in MS research remains low. Of the 758 articles included in this study, only 288 are currently available OA. This finding acts as a current benchmark against which the efforts to enhance openness and dissemination can be measured, and can serve as a reference point when assessing the impact of OA policies in this field.

The Cost of Open Access

The hybrid gold model is fraught with cost implications to both the individual researcher and institutions. The monopoly of hybrid journal ownership by the big three commercial publishers (Elsevier, Wiley, and Springer Nature) has serious implications for the openness of research findings. In this study, these 3 publishers controlled 49.7% of the journals (105/211), 49.3% of the article output (374/758), and the majority (96/105, 91.4%) of their journals followed the hybrid model.

The hybrid model had been developed as a compromise between subscription publishers and OA advocates [19]. However, it has underperformed, with less than 2% of eligible authors making work available in this manner [20]. Our study confirmed the low use of APCs, with 10% of eligible articles being made available with this option. In a recent survey of Canadian scientists, 78% felt that publishing OA was unaffordable and 86% felt that funding for OA publishing was not readily available [21]. Recent studies continue to report that lack of funds for APCs is an impediment to publishing in OA or hybrid journals [22-24].

The mode APC cost for hybrid journals was US $3000, which was 45% higher than for pure gold journals, a finding which is consistent with other recent research [25,26]. The impact of high APC costs and institutions costs has notable economic consequences in Canada, where the majority of researchers and universities are publicly funded. Academic libraries pay publishers large fees for institutional subscriptions to journals then authors pay the same publishers additional APC costs to publish their article in those same journals. In effect, there is “double dipping” of public funding, especially in the context where Canadian universities are subscribing to the “big journal deals” with all the major publishers, through the Canadian Knowledge Research Network (CRKN). In this Canadian-context study, it is also important to consider that these figures represent US dollar calculations and that exchange rate fluctuations add additional burdens to fund APCs.

Pure Gold Publication Patterns

The pure gold road to OA is not a road well-travelled by MS researchers. Pure gold journals accounted for 7.9% (60/758) of the articles and 10.4% (22/211) of the journals in this study. Although this study showed evidence of the use of pure gold journals by the researcher, the low number of articles in these venues may indicate that MS researchers feel that OA is poorly regulated, of poor quality, and lacking in peer review. This would support previous research that has found that despite awareness of OA and OA issues, faculty concerns about quality, reputation, copyright, plagiarism, and a perceived lack of peer review remained constant [27,28]. The appearance of *PLoS ONE*, ranked 11th based on article count of the 211 journals, is a promising evidence of the growing influence of gold OA.

Barriers to Green Open Access

Self-archiving is the most cost-effective method of providing OA to research findings for researchers. Local self-archiving includes depositing a postprint or preprint of the article in an IR or self-archiving in a subject repository. However, rather than offering a simplified option, publishers have created barriers to OA through unsustainable complexities involved in self-archiving policies. Of the 35 professional society or association journals represented in this study, 15 had restrictive yellow self-archiving policies. Yellow self-archiving only allows authors to post the prerefereed preprint. With peer review considered the most important criteria in journal publishing, there is potentially little value to the open communication of research of an article without peer review [29,30].

Poltronieri et al [31] also found that in their survey of journals, more than half of the publishers are still imposing yellow and white restrictions on self-archiving. Even among green journals, navigating restrictions and permissions is challenging. In addition to the version requirements of the publisher, there were additional requirements regarding embargo periods, restricting when the self-archived version could be made available. The journal embargoes were often listed as a set of complex conditions that authors would need to navigate to comply with the journal’s requirements for self-archiving.

Future Directions

The solution to resistance, either from journals or researchers, is sometimes assumed to be the requirements from funding sources or institutions [32,33]. Mandated public access from funding agencies, principally National Institutes of Health (NIH), has led to tremendous growth in the availability of biomedical literature [34]. However, the NIH, despite having an OA policy written into law 2008, found that low compliance continued to be a major issue. Researchers had indicated that a lack of time, a frustrating deposit process, and confusing journal policies were the primary reasons for lack of compliance [35]. NIH introduced a policy delaying applications and funding if publications associated with the research were not in compliance. Following this, NIH saw aggregate submissions increase from an average of 5158 articles per month in 2012 to 7931 articles per month in 2013 and 7057 in 2014 [36]. Although NIH operates in a much different context than the MS Society of Canada, the NIH experience highlights that the requirement itself is not a sufficient motivation to overcome perceived barriers to compliance. Policy enforcement is a necessary component in this process. Whereas the MS Society of Canada has implemented an OA policy, the mechanisms for enforcement are not explicitly outlined in that policy.

Whether mandated OA will move MS research to a more open dissemination environment remains to be seen. The movement to OA requires incentives for involvement [37]. In the case of researchers, for whom promotion and tenure decisions may be significantly impacted by the number of publications and the venue in which they are published, journal selection may be a critical choice. Authors select journals to publish in based on journal reputation, impact factor, and turnaround time to publication [22,28,38]. Watson [39] found that the authors are confused by the notion of OA, reluctant to participate, and confounded by the myriad of choices they are presented with when trying to publish OA.

Surveys have shown that researchers have generally favorable views of OA and its benefits for both the public and the scientific community [40,41]. However, there are challenges to making work available through both green and gold roads to

OA. Initial challenges include informing researchers about the concept of OA itself, as many scientists and researchers have no direct mechanisms in place to become more informed about these issues and may not recognize the necessity of doing so [42].

Study Limitations

As this study relied on bibliographic analysis rather than contacting researchers directly to provide a list of articles, the 758 articles used in the analysis may not have been directly produced as a result of the MS Society of Canada funded research. Any participating authorship by the 77 researchers during the time of funding by the Society was included in the study as long as the article was related to the topic of MS.

Conclusions

The prevalence of OA literature produced by the MS Society of Canada researchers has remained consistently low between 2009 and 2014. Of the 758 articles, 288 are available OA. Most OA articles were made available without evidence of APC payment. Although APCs in hybrid journals was significantly higher than those associated with pure gold journals, the use of APCs for publication was low among this group.

The recent implementation of the MS Society of Canada's OA policy may increase OA publication within this field. Future research should include prevalence of OA of newly funded research as a means of determining policy impact and effectiveness.

Acknowledgments

This work was supported by the Research Support Fund of Wilfrid Laurier University and Alberta Health Innovates.

Conflicts of Interest

None declared.

References

1. Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG. Multiple sclerosis. N Engl J Med 2000 Sep 28;343(13):938-952. [doi: 10.1056/NEJM200009283431307] [Medline: 11006371]
2. Beck CA, Metz LM, Svenson LW, Patten SB. Regional variation of multiple sclerosis prevalence in Canada. Mult Scler 2005 Oct;11(5):516-519. [Medline: 16193887]
3. Atkins HL, Bowman M, Allan D, Anstee G, Arnold DL, Bar-Or A, et al. Immunoablation and autologous haemopoietic stem-cell transplantation for aggressive multiple sclerosis: a multicentre single-group phase 2 trial. Lancet 2016 Aug 06;388(10044):576-585. [doi: 10.1016/S0140-6736(16)30169-6] [Medline: 27291994]
4. Wollin J, Dale H, Spenser N, Walsh A. What people with newly diagnosed MS (and their families and friends) need to know. Int J MS Care 2000;2(3):29-39 [FREE Full text] [doi: 10.7224/1537-2073-2.3.29]
5. Multiple Sclerosis Society of Canada. 2015. Operating Grant Policies URL: https://mssociety.ca/uploads/files/og-policies-final-en.pdf [accessed 2016-06-08] [WebCite Cache ID 6i72KRDiL]
6. Narayanaswami P, Gronseth G, Dubinsky R, Penfold-Murray R, Cox J, Bever C, et al. The Impact of Social Media on Dissemination and Implementation of Clinical Practice Guidelines: A Longitudinal Observational Study. J Med Internet Res 2015;17(8):e193 [FREE Full text] [doi: 10.2196/jmir.4414] [Medline: 26272267]
7. Colombo C, Filippini G, Synnot A, Hill S, Guglielmino R, Traversa S, et al. Development and assessment of a website presenting evidence-based information for people with multiple sclerosis: the IN-DEEP project. BMC Neurol 2016 Mar 02;16:30 [FREE Full text] [doi: 10.1186/s12883-016-0552-0] [Medline: 26934873]
8. Koschack J, Weibezahl L, Friede T, Himmel W, Makedonski P, Grabowski J. Scientific versus experiential evidence: discourse analysis of the chronic cerebrospinal venous insufficiency debate in a multiple sclerosis forum. J Med Internet Res 2015 Jul 01;17(7):e159 [FREE Full text] [doi: 10.2196/jmir.4103] [Medline: 26133525]
9. Köpke S, Solari A, Khan F, Heesen C, Giordano A. Information provision for people with multiple sclerosis. Cochrane Database Syst Rev 2014 Apr 21(4):CD008757. [doi: 10.1002/14651858.CD008757.pub2] [Medline: 24752330]
10. Lejbkowicz I, Paperna T, Stein N, Dishon S, Miller A. Internet usage by patients with multiple sclerosis: implications to participatory medicine and personalized healthcare. Mult Scler Int 2010;2010:640749 [FREE Full text] [doi: 10.1155/2010/640749] [Medline: 22096625]
11. Colombo C, Mosconi P, Confalonieri P, Baroni I, Traversa S, Hill SJ, et al. Web search behavior and information needs of people with multiple sclerosis: focus group study and analysis of online postings. Interact J Med Res 2014;3(3):e12 [FREE Full text] [doi: 10.2196/ijmr.3034] [Medline: 25093374]
12. Heesen C, Kasper J, Segal J, Köpke S, Mühlhauser I. Decisional role preferences, risk knowledge and information interests in patients with multiple sclerosis. Mult Scler 2004 Dec;10(6):643-650. [Medline: 15584489]
13. Snyder J, Adams K, Crooks VA, Whitehurst D, Vallee J. "I knew what was going to happen if I did nothing and so I was going to do something": faith, hope, and trust in the decisions of Canadians with multiple sclerosis to seek unproven interventions abroad. BMC Health Serv Res 2014 Sep 30;14:445 [FREE Full text] [doi: 10.1186/1472-6963-14-445] [Medline: 25265935]

14. Marrie RA, Salter AR, Tyry T, Fox RJ, Cutter GR. Preferred sources of health information in persons with multiple sclerosis: degree of trust and information sought. J Med Internet Res 2013;15(4):e67 [FREE Full text] [doi: 10.2196/jmir.2466] [Medline: 23635393]
15. Multiple Sclerosis Society of Canada. 2016. Research Summaries URL: https://mssociety.ca/about-our-research-program/research-we-fund [accessed 2016-06-08] [WebCite Cache ID 6i72Ws2wa]
16. Gafson AR, Giovannoni G. CCSVI-A. A call to clinicans and scientists to vocalise in an Internet age. Mult Scler Relat Disord 2014 Mar;3(2):143-146. [doi: 10.1016/j.msard.2013.10.005] [Medline: 25878001]
17. Pullman D, Zarzeczny A, Picard A. "Media, politics and science policy: MS and evidence from the CCSVI Trenches". BMC Med Ethics 2013 Feb 12;14:6 [FREE Full text] [doi: 10.1186/1472-6939-14-6] [Medline: 23402260]
18. Ghahari S, Forwell SJ. Social media representation of chronic cerebrospinal venous insufficiency intervention for multiple sclerosis. Int J MS Care 2016;18(2):49-57 [FREE Full text] [doi: 10.7224/1537-2073.2014-073] [Medline: 27134577]
19. Prosser D. From here to there: A proposed mechanism for transforming journals from closed to open access. Learn Publ 2003;16(3):163-166 [FREE Full text] [doi: 10.1087/095315103322110923]
20. Björk B. The hybrid model for open access publication of scholarly articles: A failed experiment? J Am Soc Inf Sci Technol 2012;63(8):1496-1504 [FREE Full text] [doi: 10.1002/asi.22709]
21. Björk B, Solomon D. Open access versus subscription journals: a comparison of scientific impact. BMC Med 2012;10(1):73 [FREE Full text] [doi: 10.1186/1741-7015-10-73] [Medline: 22805105]
22. Phase 5 Research. Canadian Science Publishing. 2014. Canadian Researchers' Publishing Attitudes and Behaviours URL: http://www.cdnsciencepub.com/files/PDF/CSP_ResearcherAttitudes_March14_FINAL.pdf [accessed 2016-06-08] [WebCite Cache ID 6i72geQbb]
23. Lwoga E, Questier F. Faculty adoption and usage behaviour of open access scholarly communication in health science universities. New Libr World 2014;115(3/4):116-139 [FREE Full text] [doi: 10.1108/NLW-01-2014-0006]
24. Togia A, Korobili S. ELPUB. Attitudes towards open access: a metasynthesis of the empirical literature URL: http://elpub.scix.net/data/works/att/102_elpub2014.content.pdf [accessed 2016-06-08] [WebCite Cache ID 6i730drBV]
25. Pinfield S, Salter J, Bath P. The 'total cost of publication' in a hybrid open-access environment: Institutional approaches to funding journal-article proccessing charges in combination with subscriptions. J Am Soc Inf Sci Technol 2016;67(7):1751-1766 [FREE Full text]
26. Solomon D, Björk B. Article processing charges for open access publication-the situation for research intensive universities in the USA and Canada. Peer J 2016;4:e2264 [FREE Full text] [doi: 10.7717/peerj.2264] [Medline: 27547569]
27. Xia J. A longitudinal study of scholars attitudes and behaviors toward open-access journal publishing. J Am Soc Inf Sci Technol 2010;61(3):615-624 [FREE Full text] [doi: 10.1002/asi.21283]
28. Lwoga ET, Questier F. Open access behaviours and perceptions of health sciences faculty and roles of information professionals. Health Info Libr J 2015 Mar;32(1):37-49. [doi: 10.1111/hir.12094] [Medline: 25603838]
29. Moore G. University of Toronto. 2011. Survey of University of Toronto faculty awareness, attitudes and practices regarding scholarly communication: A preliminary report URL: https://tspace.library.utoronto.ca/handle/1807/26446 [accessed 2016-06-08] [WebCite Cache ID 6i73r6RhV]
30. Austin A, Heffernan M, David N. Open Access to Knowledge (OAK) Law Project. 2008. Academic authorship, publishing agreements and open access: Survey results URL: http://eprints.qut.edu.au/13623/1/13623_3.pdf [accessed 2016-06-08] [WebCite Cache ID 6i73wNrro]
31. Poltronieri E, Bravo E, Camerini T, Ferri M, Rizzo R, Solimini R, et al. Where on earth to publish? A sample survey comparing traditional and open access publishing in the oncological field. J Exp Clin Cancer Res 2013;32:4 [FREE Full text] [doi: 10.1186/1756-9966-32-4] [Medline: 23339627]
32. Harnad S. Open access to research: Changing researcher behavior through university and funder mandates. JeDEM 2011;3(1):33-41 [FREE Full text]
33. Pinfield S. A mandate to self archive? The role of open access institutional repositories. Serials: The Journal for the Serials Community 2005 Mar 1;18(1):30-34. [doi: 10.1629/1830]
34. Morrison H. The Imaginery Journal of Poetic Economics. 2015. Dramatic Growth of Open Access URL: http://poeticeconomics.blogspot.ca/2015/06/dramatic-growth-of-open-access-june-30.html [accessed 2016-06-08] [WebCite Cache ID 6i747hwtD]
35. Charbonneau DH, McGlone J. Faculty experiences with the National Institutes of Health (NIH) public access policy, compliance issues, and copyright practices. J Med Libr Assoc 2013 Jan;101(1):21-25 [FREE Full text] [doi: 10.3163/1536-5050.101.1.004] [Medline: 23405043]
36. National Institutes of Health. 2015. Monthly Aggregate Submission Statistics URL: http://www.nihms.nih.gov/stats/ [accessed 2016-06-08] [WebCite Cache ID 6i74EYtzy]
37. Parks R. The Faustian grip of academic publishing. J Econ Methodol 2010;9:317-335 [FREE Full text] [doi: 10.1080/1350178022000015122]
38. Kuchma I. Electronic Information for Libraries. 2011. Results of the SOAP survey: A preliminary overview of the situation in EIFL partner countries URL: http://www.eifl.net/system/files/resources/201408/soap_survey_results_eifl.pdf [accessed 2016-06-08] [WebCite Cache ID 6i74HtBA3]

39. Watson M. When will 'open science' become simply 'science'? Genome Biol 2015;16:101 [FREE Full text] [doi: 10.1186/s13059-015-0669-2] [Medline: 25986601]
40. Dallmeier-Tiessen S, Darby R, Goerner B, Hyppoelae J, Igo-Kemenes P, Kahn D. Cornell University Library. 2011. Highlights from the SOAP project survey. What scientists think about open access publishing URL: http://arxiv.org/abs/1101.5260 [accessed 2016-06-08] [WebCite Cache ID 6i72Zxg8p]
41. Fry J, Probets S, Creaser C, Greenwood H, Spezi V, White S. Loughborough University Institutional Repository. 2011. PEER behavioural research: authors and users vis-a-vis journals and repositories. D4.2 final report URL: https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/12967/3/PEER_D4_final_report_29SEPT11.pdf [accessed 2016-06-08] [WebCite Cache ID 6i72btafm]
42. Moorhead LL, Holzmeyer C, Maggio LA, Steinberg RM, Willinsky J. In an age of open access to research policies: Physician and public health NGO staff research use and policy awareness. PLoS One 2015;10(7):e0129708 [FREE Full text] [doi: 10.1371/journal.pone.0129708] [Medline: 26200794]

Abbreviations

APC: article processing charge
CRKN: Canadian Knowledge Research Network
IR: Institutional Repository
MS: multiple sclerosis
NIH: National Institutes of Health
OA: open access

Edited by G Eysenbach; submitted 21.06.16; peer-reviewed by J Till PhD, H Morrison; comments to author 17.11.16; revised version received 15.12.16; accepted 31.01.17; published 27.02.17.

Please cite as:
Bakker C, Stephenson C, Stephenson E, Chaves D
Public Funding and Open Access to Research: A Review of Canadian Multiple Sclerosis Research
J Med Internet Res 2017;19(2):e52
URL: http://www.jmir.org/2017/2/e52/
doi:10.2196/jmir.6250
PMID:28242594

Original Paper

Issues for eHealth in Psychiatry: Results of an Expert Survey

Jennifer Nicholas[1,2], BA, BSc (Hons); Kit Huckvale[1], MD, PhD; Mark Erik Larsen[1], DPhil; Ashna Basu[2]; Philip J Batterham[3], PhD; Frances Shaw[1], PhD; Shahbaz Sendi[4], MD, PhD

[1]Black Dog Institute, University of New South Wales Australia, Sydney, Australia

[2]School of Psychiatry, University of New South Wales Australia, Sydney, Australia

[3]Centre for Mental Health Research, Australian National University, Canberra, Australia

[4]Affective Disorders Research Group, Center for Affective Disorders, Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

Corresponding Author:
Jennifer Nicholas, BA, BSc (Hons)
Black Dog Institute
University of New South Wales Australia
Hospital Rd
Prince of Wales Hospital, Randwick
Sydney,
Australia
Phone: 61 9382 9260
Fax: 61 9382 8208
Email: j.nicholas@blackdog.org.au

Abstract

Background: Technology has changed the landscape in which psychiatry operates. Effective, evidence-based treatments for mental health care are now available at the fingertips of anyone with Internet access. However, technological solutions for mental health are not necessarily sought by consumers nor recommended by clinicians.

Objective: The objectives of this study are to identify and discuss the barriers to introducing eHealth technology-supported interventions within mental health.

Methods: An interactive polling tool was used to ask “In this brave new world, what are the key issues that need to be addressed to improve mental health (using technology)?” Respondents were the multidisciplinary attendees of the “Humans and Machines: A Quest for Better Mental Health” conference, held in Sydney, Australia, in 2016. Responses were categorized into 10 key issues using team-based qualitative analysis.

Results: A total of 155 responses to the question were received from 66 audience members. Responses were categorized into 10 issues and ordered by importance: access to care, integration and collaboration, education and awareness, mental health stigma, data privacy, trust, understanding and assessment of mental health, government and policy, optimal design, and engagement. In this paper, each of the 10 issues are outlined, and potential solutions are discussed. Many of the issues were interrelated, having implications for other key areas identified.

Conclusions: As many of the issues identified directly related to barriers to care, priority should be given to addressing these issues that are common across mental health delivery. Despite new challenges raised by technology, technology-supported mental health interventions represent a tremendous opportunity to address in a timely way these major concerns and improve the receipt of effective, evidence-based therapy by those in need.

(J Med Internet Res 2017;19(2):e55) doi:10.2196/jmir.6957

KEYWORDS

eHealth; mental health; technology adoption

Introduction

The enormous personal, social, and financial burden caused by mental health problems is increasingly acknowledged. Mental health conditions are the leading cause of years lost to disability globally and account for 8.9% of global disability-adjusted life years [1]. A higher risk of suicide and increased medical comorbidity means that individuals with serious mental illness die up to 32 years earlier than the general population [2].

With the emergence of digital technologies, the landscape in which psychiatry operates has changed. There are now an expanded array of tools and resources at our disposal. Electronic health (eHealth) refers to a range of services that use information and communication technologies to improve human health. Examples include programs that deliver cognitive behavioral therapy to anonymous users over the Internet [3], remote assessment and treatment of patients using telecommunications and the Internet [4], and novel treatment interventions via virtual reality or serious games [5]. Recently, eHealth has expanded to include mobile health (mHealth), which aims to harness consumer-facing technologies such as smartphones and wearable devices to support health care.

In addition to direct delivery of services, eHealth technology can assist the administration and coordination of mental health care. For example, in the United Kingdom, the National Health Service (NHS) is integrating technology into primary care to allow patients to register with a general practitioner (GP), access health care records, and receive medical advice and information via their computer, smartphone, or tablet [6]. Allowing patients immediate access to private medical records via the Internet empowers consumers and may facilitate consistency of care.

However, while the promise of technology in mental health is high, delivery and uptake remains low. In the United States, only 20% of organizations involved in provision of behavioral support have adopted electronic records, compared with 60% of primary care organizations [7]. In Australia, an estimated 600,000 individuals with mild-moderate mental health disorders are potentially suitable for eHealth services, while only 25% currently receive any mental health care [8]. As yet, health system integration and the management, engagement, and prevention of mental health on a population scale remain unrealized possibilities. This is despite acceleration in consumer use of technology, with rapid adoption of platforms such as smartphones and new services for banking, commerce, travel, and social interaction.

There is a need to assess the reasons for this disparity. Is it just a matter of time [9], or are there barriers within mental health that will impede the deployment of technological solutions? Are the issues technological or do they arise from other sources? It is not clear what these problems, barriers, or concerns might be. Identification and discussion of possible issues may inform strategies to ensure the potential of technology for mental health is realized.

Taking advantage of a concentration of expertise drawn from across the mental health sector in Australia, we surveyed the audience of the “Humans and Machines: A Quest for Better Mental Health” conference about the key issues that need to be addressed to improve mental health using eHealth technologies. Based on audience responses, this paper aims to outline the perceived key issues and suggest ways to overcome these barriers.

Methods

Data Collection

Data were collected via a 1-question, cross-sectional, interactive survey using a convenience sample. In September 2016, the Black Dog Institute and University of New South Wales (UNSW) Australia hosted the “Humans and Machines: A Quest for Better Mental Health” conference in Sydney, Australia. The aim of the conference was to explore the interface between science, eHealth technologies, and human health and whether a physical face-to-face presence is required to provide quality mental health care. The survey question “In this brave new world, what are the key issues that need to be addressed to improve mental health (using technology)?” was presented to attendees at the end of the first session via the Poll Everywhere interactive data collection tool [10]. Attendees anonymously provided their free-text responses using the browser on their mobile devices. No limit was placed on the number of responses submitted by each individual.

Data Analysis

Audience responses to the question were organized into key issue areas following guidelines for rigorous team-based approaches to decision making [11]. Two authors (JN and AB) independently generated 10 data-driven issue areas from the audience responses. Identification of issue areas was mostly inductive, allowing the analysis to be flexible and theory independent but guided by the research question [12]. Between the 2 coders, 11 unique issues were identified, of which 9 were identified by both coders. A third party (KB) with substantial experience with qualitative research resolved the difference in issue identification to provide a final list of 10 issues and any differences in individual response categorization in consultation with JN and AB. Issue importance was proxied by calculating the percentage of respondents that nominated each of the 10 issues identified.

Results

A total of 94 individuals attended the “Humans and Machines: A Quest for Better Mental Health” conference. Speakers and audience members included a broad representation of senior staff spanning eHealth research, mental health professionals, health service providers, philanthropic organizations, and the health and technology industry. A total of 155 unique responses to the question were submitted by 66 audience members, from which 10 key issues that need to be addressed to improve mental health using technology were identified. The 10 issues identified were interrelated and are displayed in order of importance in Table 1.

Table 1. The 10 issues identified that need to be addressed to improve mental health (using technology) ranked in order of importance.

Issue identified	n (%)
Access to care	24 (36)
Integration and collaboration	15 (23)
Education and awareness	13 (20)
Mental health stigma	13 (20)
Data privacy	12 (18)
Trust	11 (17)
Understanding and assessment of mental health	11 (17)
Government and policy	10 (15)
Optimal design	9 (14)
Engagement	8 (12)

Discussion

Principal Findings

By combining the perspectives of a wide range of stakeholders drawn from a recent technology-focused conference, this opportunistic survey sought to provide a contemporary overview of shared priority issues that need to be considered if the potential of eHealth is to be realized in mental health. Rather than prespecify a technology, condition, or policy focus, the survey was intended to solicit the broadest range of opinions possible in order to understand the extent to which prevalent issues are technology-specific or, rather, represent an extension of known challenges in mental health generally. Validating this approach, of the identified issues, half (n=5) reflected pragmatic concerns of access, understanding, and attitudes to mental illness that extend beyond technology to stand as common barriers to improved mental health care. Technology cannot escape these issues. Yet there should also be considerable optimism in the potential for eHealth technologies to offer novel, substantive strategies to tackle these barriers to care.

Reflecting these dual notions of challenge and opportunity, the following discussion attempts to highlight how eHealth technology is shaped by and holds the potential to shape each identified issue. In addition to the service delivery themes identified above, the remaining issues fell into 2 further categories: structural issues surrounding mental health policy and services (n=2) and technology-specific issues (n=3).

Issues Affecting Mental Health Service Delivery

Access to Care

Participants overwhelmingly highlighted the need to improve the timely access of mental health care by those in need. As well as general improvements in care access, reducing social inequalities in accessing mental health care was emphasized. Despite the range of effective strategies currently available to treat mental health conditions, too few individuals seek help. Projections indicate that improving service access among the two-thirds of Australians with a mental health disorder not receiving care would result in a 23% reduction in the burden of common mental disorders [13,14]. Addressing access to care involves consideration of barriers to care, many of which were identified by participants as issues requiring attention to improve mental health, including stigma, education, and health system integration and policy considerations. Additional proposed barriers to help seeking include a desire to handle the problem without outside help, distrust of mental health services [15], and concerns regarding cost, transport, time, and convenience [16].

Given the high rate of Internet access [17] and growing ownership of mobile devices [18] in both developed and developing settings, technology-supported interventions can address many of these identified barriers. Effective eHealth interventions are available in the form of unguided self-help [3,19] and can be accessed anonymously, minimizing the possibility of stigma. Such interventions provide around-the-clock access to evidence-based treatments, allowing timely access in response to symptoms and maximizing consumer convenience. eHealth interventions are, arguably, also able to provide more equitable access to health services, given their ability to be accessed remotely and at minimal to no cost to the user [20]. However, while technology may seem to address these barriers to help seeking with the potential to improve access to care, formal evaluation of the ability of technology-supported interventions to engage those otherwise not accessing care is required. Other factors that influence access to care, including trust discussed later, require consideration throughout design and development.

In aiming to address the issue of care access, technology-supported interventions must also consider scalability. Much has been made of the possibility for eHealth to have large-scale, population-based, system-wide implications for mental health [8]. However, populations are not homogeneous in terms of need, interest, or access to technology, all of which will mediate intervention success at scale.

Moreover, technology may itself generate inequalities in health service delivery and access. For example, technology-supported services often target younger adults based on their presumed affinity for technology [21]. Yet aptitude for and uptake of technological interventions may vary within generations as much as between. From divergent rates of device ownership [22] to the challenges of digital literacy [23], the same sociodemographic forces that shape health inequalities appear also to shape eHealth access. Beyond a need to design for diversity in technological experience, these findings recommend a particularly high bar in terms of demonstrating equity of access for any eHealth strategies that aim to replace existing mental health services.

Education and Awareness

Participants identified a need for further education in 2 mental health domains, mental health literacy and availability of eHealth interventions. Poor mental health literacy, defined as the lack of "knowledge and beliefs about mental disorders, which aid their recognition, management, or prevention" [15], has been identified as a key barrier to help seeking. Technology has the potential to be an important tool in educating the general population about mental health disorders with an aim to improve mental health literacy. The Internet has been successfully used to increase public education and awareness of mental and other health conditions [24,25], with resulting increases in intentions to seek help [24,25]. Further, the Internet is increasingly used as source of information about mental health, particularly among those with mental health problems [26].

Technology-supported mental health resources must consider how to disseminate information to the intended end-users, given eHealth interventions are currently not necessarily sought by consumers nor recommended by clinicians. Efforts for integration of eHealth into existing health care systems, discussed below, will increase clinician awareness of technology-supported interventions. Technology can support this effort—for example, the eMental Health in Practice (eMHPrac) initiative uses online continuing professional development–accredited learning modules supported by webinars, forums, and blogs to teach GPs, psychiatrists, and allied health professionals about eHealth, its efficacy, and role in routine care.

Similarly, there is a need to educate the public about eHealth options, which will support their role in clinical care, as patients will inquire about eHealth even when not suggested by their clinician. To this end, many online portals such as Beacon [27] and mindhealthconnect [28] seek to inform consumers about eHealth options and guide them to the most relevant evidence-based resources. However, given the instrumental role of Facebook, Twitter, and YouTube in the unprecedented success of the Ice Bucket Challenge in raising awareness (and funds) for amyotrophic lateral sclerosis in 2014 [29], the potential for technology, in particular social media, to increase mental health literacy and eHealth awareness is not currently fully realized.

Mental Health Stigma

A need to reduce the stigma associated with mental health conditions was evident in participant responses. Negative attitudes toward people (the self or others) with mental illness has been shown to be associated with lower intentions to seek help [30,31] and delayed or diminished recovery [32,33]. In an effort to reduce fear of stigma, online therapy today is often provided anonymously, minimizing the possibility that individuals will be identified as service users and avoiding the need to label individuals as having a mental health disorder. Yet while affordances offered by technologies such as anonymity and remote interaction may help to minimize stigma exposure, prevalent negative attitudes also slow technology adoption by reducing awareness of these new options, slowing help-seeking, and focusing technology discussion on physical health.

More radical is the potential for interventions that seek to modify stigmatizing attitudes directly, potentially at population-scale. In support of this potential, a Canadian media and social media campaign raised mental health awareness and reduced mental health stigma among young adults [34]. Internet-based interventions that provide either evidence-based therapy or psychoeducation have also been shown to reduce negative perceptions of mental illness and may increase the likelihood that individuals will seek professional help [25,35]. Further research should aim to identify other avenues for using technology to reduce stigma. For example, research has found that contact interventions (involving interpersonal contact with members of the stigmatized group) lead to greater reductions in stigma than providing psychoeducation alone [36]. Although video-based contact interventions may have less impact than face-to-face contact, the use of interactive games and other immersive technological experiences may provide an effective new avenue to combating the stigma of mental illness. The potential for social media to be harnessed as an advocacy tool to reduce stigma at a population level should also continue to be explored.

Understanding and Assessment of Mental Health

Shortcomings in the understanding of mental health conditions were highlighted by participants, including established difficulties in the field relating to assessment and classification, psychosocial determinants, and prevention [37]. Despite great advances in genetic and biological medicine, translation of these developments to the understanding and treatment of mental health disorders remains incomplete [2]. Further, individual differences in symptom presentation and disease courses, coupled with the subjective nature of mental health conditions, present challenges to the development of technology-supported resources for the understanding, treatment, and prevention of these disorders.

At the intersection of big data and psychiatry, consumer-facing technologies promise access to a vast array of personal and behavioral data. Given large enough datasets, previously hidden correlations—digital biomarkers—may yet emerge with the potential to better predict outcomes and further our understanding of mental health conditions. Current apps and wearables have the ability to passively collect data about activity (from Global Positioning System sensors and accelerometers), social connectedness (from Bluetooth connectivity, social media activity, and call and text logs), sleep/wake cycles (through light sensors and screen activation), and voice tone (from microphones).

However, both passively collected data and digital mental health resources must first be linked to clinically important outcomes and assessed using validated measures that accurately account for the continuum of mental health problems in the community.

In mHealth, consensus regarding which patient-reported outcome measures are required to meaningfully assess app efficacy is needed [38].

Recent work has correlated passively collected objective data with clinically rated symptoms of depression and mania in bipolar disorder [39]. Further work with digital biomarkers will explore if data can predict changes in affective states, guide relapse prevention and clinical intervention, and ultimately inform the field about development [40], prodromes, and subclinical states [39]. Passively collected and user inputted data can also inform consumers and increase insight into their mental health, which is important for self-management.

Technology can also be used to assess risk. Digital footprints left by individuals' online presence, in particular their social media use, have been used to assess risk of depression and suicide, highlighting the extension of technology to mental illness prevention [41].

Engagement

Participants emphasized prevalent challenges of engagement with mental health therapy. Nearly half of patients with a major depressive disorder drop out of therapy within 12 months [42]. Problems of engagement disproportionately affect young people [43] and those from minority groups, who are at least 40% more likely to discontinue treatment for a mood disorder, anxiety, or depression prematurely [44].

Technology-supported interventions have similar (or greater for open access resources) difficulties with engagement [45]. Although partly reflecting issues of stigma, confidentiality, and trust (identified as separate issues in this discussion), poor engagement encompasses additional factors that reduce the likelihood of meaningful initiation, participation in, and completion of therapy once enlisted in care. In addition to accidental causes of missed treatment, such as forgotten appointments or technical issues, mental health–specific factors include variable perceptions of treatment utility among patients and carers [46], prevalent delays around treatment [47,48], the alignment of available services with personal conceptions of mental health, and preferences for different styles and modes of therapy [46,49]. Poor engagement also extends to participation in mental health service design [50] and research [51].

However, technology-supported care has the promise of addressing the common mismatch between patient expectations and service capabilities, made possible by the ability to tailor content, motivational elements, and reminders to provide personalized therapy. Self-guided therapies, available through personal devices, can be initiated without delay in response to changes in condition state and pursued at times convenient to patients. Technology-based care can simultaneously support a spectrum of peer-, clinician-, community-, and agent-based interactions that offer genuine social support for some while guaranteeing autonomous self-care for others [3,52]. More broadly, technology platforms that integrate with social media have the potential to make positive contributions to discourse about mental health by sharing information about treatment and outcomes. Further, the enhanced ability to collect unobtrusive feedback and treatment participation data promises to accelerate future developments in the field.

Structural Issues Surrounding Mental Health Policy and Services

Integration and Collaboration

Participants identified deficiencies in coordination of care and a perceived lack of interdisciplinary collaboration within the health care system as a challenge to technology-supported care. In particular, participants acknowledged the risk that new eHealth technologies perpetuate—or even extend—known challenges of service fragmentation that threaten continuity of care [53,54]. Indeed, the slow adoption of electronic health records in behavioral health settings continues to frustrate attempts to link services [7]. Many technology-supported interventions today sit independent of existing health care systems [8]. In addition to the potential complexity for decision making and patient choice arising from these new services, technology development models that do not emphasize clinical stakeholder involvement risk creating services that are a poor fit with referral and care pathways or back-office requirements, such as audit and billing.

Despite this, technology has the potential to address these challenges through better information sharing, better use of information contained in health records, and more effective communication between professionals, patients, and carers. The success of this cohesive picture will hinge on the successful and timely introduction of technology into existing systems of care.

One promising integration approach is the stepped care model in which technology-supported interventions are incrementally introduced as part of a continuum of therapies of differing intensity targeting a specific condition [8]. Such interventions have proven success in treating mild-moderate anxiety and depression [19] and release traditional resources to serve individuals with more severe symptoms [8]. It is projected that stepped care will increase quality of care for consumers and lower mental health costs by providing cost-effective care to those with mild-moderate disorders while reducing the burden on face-to-face services and increasing workforce participation.

Systemic change will require not only government support but, critically, buy-in from organizations involved in mental health delivery and leadership from clinical champions [55]. Increased awareness of technological mental health interventions among consumers and clinicians aided by mental health professional education and training will be important [8]. Further, shared information technology infrastructure and successful deployment of electronic health records will be necessary to ensure continuity of care from technology-supported to in-person services [8]. Substantive interdisciplinary collaboration that incorporates consumer perspectives will also become increasingly important as the dependencies between services and systems grows [56].

Government and Policy

For the field to flourish, participants suggested governments will need to develop frameworks and policies to encourage innovation and technology within health services. Recent findings indicate that government-facilitated access to electronic medical records through patient portals increases consumer involvement in health care and improves health outcomes [57,58]. The Australian government has recognized the potential and cost effectiveness of technology in mental health care service delivery [59], supporting the e-mental health record and the eMHPrac initiative designed to promote online mental health resources in primary care [60]. Dedicated centers of excellence for eHealth research are also supported, as well as organizations providing services directly to the public which receive approximately 275,000 combined unique website visits each month [8].

In addition to mentioning an enabling role for policy makers, participants also highlighted the pressing need for timely governance of emerging technologies characterized by a rapid pace of change. For example, although regulation concerning the development of health apps now exists in some health economies, notably in the United States [61], the scope of these guidelines is restricted to diagnostic and therapeutic categories that commonly exclude mental health. This patchy regulatory coverage allows anyone to develop and deploy an app for mental health through commercial app stores without an evidence base. Clinical assessments of mental health apps have highlighted a lack of evidence-based content and minimal demonstrations of efficacy [62,63]. While this may not affect consumer uptake, clinicians are understandably wary of recommending apps to support treatment in the absence of quality guarantees, slowing integration into care. Beyond apps, the emerging scope for technology-supported population-scale digital mental health prevention and health promotion campaigns will open up new issues around data governance and ethics that lie outside existing governance frameworks. Market-initiated solutions such as clearing houses, development guidelines, and quality checklists may have a role [64], but effective regulation of medical technologies has historically relied on government intervention.

Technology-Specific Issues

Data Privacy

Participants highlighted the need to safeguard the privacy of identifiable, sensitive health information collected by eHealth services. Secure data storage and the choice to remain anonymous were considered necessities when dealing with health data. Further, given the ability to passively collect an unprecedented diversity and volume of personal and behavioral data through smartphone apps, responses emphasized informed, user-controlled data collection.

Unfortunately, there is often considerable opacity regarding data collection and processing in technological interventions, illuminated only by privacy policies that are often long and difficult to understand. The availability of a privacy policy is not, however, a comprehensive solution for understanding the privacy implications of a technological intervention. A study of app privacy found the majority of policies did not actually focus on the app concerned [65]. Furthermore, privacy policies do not necessarily reflect what happens “inside the black box,” as 78% of accredited apps in the now defunct NHS Health Apps Library uploaded data which had not been disclosed to users [66]. Although these limitations may not factor in user decisions on whether or not to use an eHealth platform, they nevertheless reflect noncompliance with relevant privacy regulations, which hinder the possibility of ethical, informed decisions regarding the use of specific platforms. To assist consumers, as well as ensuring the enforcement of privacy regulations, the provision of simplified user-friendly privacy information has been proposed, akin to the recent overhaul in the presentation of nutritional information [67].

Trust

Reflecting wider conceptions of distrust of mental health services as a barrier to care [15], participants identified multiple dimensions to the concept of trust in the use of technology for mental health, including ethical data collection and analysis and the need for its responsible use. The costs of breaching trust were considered catastrophic. Organizations involved in data collection and analysis or delivering eHealth interventions need to have the trust of users. Further, users, including mental health professionals, also need to be able to trust the technologies. Medical practitioners have been found to “see data as costs, risks, and liabilities” [68].

Similarly, it is important to consumers to trust that data will be used for the public good [69]. Context is critical for these perceptions of trust, with considerations around why and how data will be used contributing to whether or not data feels right or feels wrong [70]. Therefore, transparency in data collection purposes, access, and uses should be emphasized.

There must also be trust that the technologies and techniques used for data analysis are secure and effective and will provide accurate identification of mental health symptoms or risk for mental health problems. Kennedy [71], drawing on the work of Theodore Porter, argues that quantification apparently reduces the need for interpersonal trust through the appearance of objectivity in data analysis techniques but that reliance on numbers can increase distrust if errors are made. This highlights the importance of evidence and accuracy in the use of eHealth interventions and data analysis in mental health. Therefore, it is paramount that the evidence-base for the wide range of technological interventions in mental health is developed, as is the case for online interventions for common mental health disorders [3,19]. Further, certification or accreditation for eHealth and mHealth programs could enhance consumer trust of these resources.

Optimal Design

Participants highlighted uncertainties about the optimal design of technology-supported mental health interventions. Some reflected longstanding thematic concerns for eHealth applicable not only to mental health, such as how best to translate therapeutic principles to a technology-based medium while retaining clinical effectiveness, how to identify patient groups most likely to benefit from technology-supported care, and how to appropriately tailor both platform (whether Web, app, or social media) and design to ensure usability and acceptability among target users [72,73]. While a perceived benefit of eHealth interventions is their ability to address varied experience and personal risk factors as well as cultural norms in diverse populations, this increases the complexity and cost of design.

Lessons learned in other disciplines may guide design considerations. Recent work in health promotion has highlighted the potential to design complex interventions for behavior change using discrete building blocks, contributing to theory-building, and maximizing likely effectiveness [74]. Strategies such as user-centered and participatory design, which emphasize substantive involvement of target users throughout intervention development, can simultaneously refine intervention focus while eliminating potential usability barriers [56]. New evaluation strategies, which are better suited to both iterative improvements in intervention designs and the fast past of technology change, will also be needed [75].

However, there is work to be done to optimize these techniques for the design of mental health interventions. For example, any theory seeking to maximize user interaction, whether through motivational elements or gameplay, must also be compatible with psychological theory guiding therapy. As a result, effective design requires not only technical proficiency from software developers, but commensurate skills among clinical staff to understand the conceptual basis behind concepts such as serious games and translate these in ways that achieve specific desired outcomes such as improved adherence [44] and are compatible with evaluation [76].

Limitations

The limitations of this paper require acknowledgment. The survey was conducted using a convenience sample of attendees at a conference convened to discuss the interface between science, technology, and human health and the potential role of technology in providing quality mental health care. Given a convenience sample and a common interest in technology, attendee perspectives around the challenges of mental eHealth may not be fully representative, particularly of stakeholders who have made a principled choice *not* to use eHealth technologies. However, attendees were also experts in eHealth with research and organizational roles where a balanced understanding of the issues could be reasonably expected.

A further limitation is that consumers were not represented among conference attendees, limiting the issues identified to those important from a service provision perspective. It is widely acknowledged that consumers and service providers traditionally hold different views on mental health care challenges and priorities [77], and thus an important and varied perspective is not represented in these results. Future research should address this gap and aim to understand the issues in the introduction of technology in mental health service delivery perceived by consumers.

Finally, the frequency-of-elicitation method used to assess relative importance may be an imperfect proxy for participant views. Factors other than importance that may have influenced participant submissions include accidental omission and perceptions that a topic might have already have been submitted by others. It is therefore possible that, given a forced choice method, a different ranking would have emerged. As a result, while the consistency of themes that emerged (despite the diversity of participant backgrounds) strengthens the convergent validity of the issues considered as a set, the rank order should be interpreted with caution.

Conclusions

Current mental health service provision has failed to engage a large number of those in need. Many of the issues identified by attendees of the “Humans and Machines: A Quest for Better Mental Health” conference directly relate to barriers to care, including access to care, stigma, education, engagement, integration, and government and policy. Priority should therefore be given to addressing these issues that are common across mental health delivery. Despite new challenges, technology-supported mental health interventions represent a tremendous opportunity to overcome these issues but only if they are actively considered during design and development.

Several studies have shown that the use of technology in mental health care is acceptable and at times preferable to consumers, with convenience, cost, and anonymity listed among its advantages [78]. Indeed, in some cases, the use of technology for mental health care has been largely consumer-driven (e.g., the use of apps).

However, to fulfill this potential, an integrated, coordinated approach is needed to establish a role for eHealth services within existing health care systems and increase awareness of these services among consumers and clinicians. This requires commitment from all stakeholders, including research, clinical practice, regulators, and governments to support the role of technology in mental health. Research has established the effectiveness of a range of e-mental health services, but an emphasis on implementation science is essential to ensuring the successful scaling of digital health interventions. Clinician awareness and training programs are vital to inform and support the role of eHealth in routine practice and to guarantee consumers are directed to appropriate technological interventions. Critically, while government recognition of the benefits of eHealth within the health care system is important to the delivery of eHealth programs, this recognition may be meaningless without sustained funding to maintain eHealth services and for continuing development to ensure that health interventions keep pace with emerging technology.

Without a concerted effort to translate research into policy and practice to address the barriers described here, the adoption of technology into mental health care will inevitably be slowed. However, the greater failure will be to miss the potential, offered by technology, to address in a timely way major concerns of access, stigma, and engagement that stand as active barriers to participation in mental health care and the receipt of effective, evidence-based therapy by those most in need.

Acknowledgments

Authors wish to thank UNSW Australia and The University of Sydney for convening the “Humans and Machines: A Quest for Better Mental Health” conference. We also thank the audience members for their responses regarding the key issues that need to be addressed to improve mental health and Professor Katherine Boydell for acting as the independent third party for the qualitative analysis. This study is supported by a grant from the National Health and Medical Research Council (NHMRC) and forms part of research conducted by the NHMRC Centre for Research Excellence in Suicide Prevention (CRESP; APP1042580). Additional support for the conference was provided by UNSW Brain Sciences. JN is supported by an Australian Postgraduate Award, ML is supported by a Society of Mental Health Research 2015 Early Career Research Award, and PJB is supported by NHMRC Fellowship 1083311.

Conflicts of Interest

None declared.

References

1. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61689-4] [Medline: 23245608]
2. Insel TR. Translating scientific opportunity into public health impact: a strategic plan for research on mental illness. Arch Gen Psychiatry 2009 Feb;66(2):128-133. [doi: 10.1001/archgenpsychiatry.2008.540] [Medline: 19188534]
3. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the Internet: randomised controlled trial. BMJ 2004 Jan 31;328(7434):265 [FREE Full text] [doi: 10.1136/bmj.37945.566632.EE] [Medline: 14742346]
4. Monnier J, Knapp RG, Frueh BC. Recent advances in telepsychiatry: an updated review. Psychiatr Serv 2003 Dec;54(12):1604-1609. [doi: 10.1176/appi.ps.54.12.1604] [Medline: 14645799]
5. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Aug;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
6. National Health Service. NHS general practitioners (GPs) services: GP online services URL: http://www.nhs.uk/NHSEngland/AboutNHSservices/doctors/Pages/gp-online-services.aspx [accessed 2016-11-21] [WebCite Cache ID 6mAtMIwaA]
7. RTI International Center for the Advancement of Health Information Technology. Behavioral health roundtable: using information technology to integrate behavioral health and primary care URL: https://www.healthit.gov/sites/default/files/bh-roundtable-findings-report_0.pdf [accessed 2016-11-21] [WebCite Cache ID 6mArI0zUD]
8. Christensen EM, Proudfoot J, Woodward A, Hosie A, Klein B, Morgan C. E-mental health services in Australia 2014: current and future. URL: https://emhalliance.fedehealth.org.au/wp-content/uploads/sites/42/2014/10/e-Mental-Health-in-Australia-2014.pdf [accessed 2016-11-11] [WebCite Cache ID 6lvqMQcmN]
9. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011 Dec;104(12):510-520 [FREE Full text] [doi: 10.1258/jrsm.2011.110180] [Medline: 22179294]
10. Poll Everywhere. URL: https://www.polleverywhere.com/about [accessed 2016-11-11] [WebCite Cache ID 6lvp3oz5q]
11. Erlandson D, Harris E, Skipper B, Allen S. Doing naturalistic inquiry: a guide to methods. Newbury Park, CA: Sage Publications; 1993.
12. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 8;320(7227):114-116 [FREE Full text] [Medline: 10625273]
13. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. Br J Psychiatry 2004 Jun;184:526-533 [FREE Full text] [Medline: 15172947]
14. Australian Bureau of Statistics. National Survey of Mental Health and Wellbeing URL: http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/4326.0Main%20Features12007?opendocument&tabname=Summary&prodno [accessed 2016-11-21] [WebCite Cache ID 6lpna5cpm]
15. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. “Mental health literacy”: a survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. Med J Aust 1997 Feb 17;166(4):182-186. [Medline: 9066546]
16. Mojtabai R. Unmet need for treatment of major depression in the United States. Psychiatr Serv 2009 Mar;60(3):297-305. [doi: 10.1176/appi.ps.60.3.297] [Medline: 19252041]
17. ICT Facts & Figures. Geneva: International Telecommunications Union URL: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf [accessed 2016-11-21] [WebCite Cache ID 6iE2i7VE3]
18. Ericsson Mobility Report: On the pulse of the networked society. URL: http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-feb-2015-interim.pdf [accessed 2016-11-21] [WebCite Cache ID 6Y4KpDV9G]

19. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
20. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. J Ment Health 2011 Dec;20(6):509-524. [doi: 10.3109/09638237.2011.608746] [Medline: 21988230]
21. Brown C, Czerniewicz L. Debunking the “digital native”: beyond digital apartheid, towards digital democracy. J Comput Assist Lear 2010;26(5):357-369. [doi: 10.1111/j.1365-2729.2010.00369.x]
22. U.S. smartphone use in 2015. Chapter one: a portrait of smartphone ownership. Washington: Pew Internet and American Life Project URL: http://www.pewinternet.org/2015/04/01/chapter-one-a-portrait-of-smartphone-ownership/ [accessed 2017-01-29] [WebCite Cache ID 6nt6mBY6i]
23. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 04;18(10):e264 [FREE Full text] [doi: 10.2196/jmir.6349] [Medline: 27702738]
24. Brijnath B, Protheroe J, Mahtani KR, Antoniades J. Do Web-based mental health literacy interventions improve the mental health literacy of adult consumers? Results from a systematic review. J Med Internet Res 2016 Jun 20;18(6):e165 [FREE Full text] [doi: 10.2196/jmir.5463] [Medline: 27323907]
25. Taylor-Rodgers E, Batterham PJ. Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: randomised controlled trial. J Affect Disord 2014 Oct 15;168:65-71. [doi: 10.1016/j.jad.2014.06.047] [Medline: 25038293]
26. Powell J, Clarke A. Internet information-seeking in mental health: population survey. Br J Psychiatry 2006 Sep;189:273-277 [FREE Full text] [doi: 10.1192/bjp.bp.105.017319] [Medline: 16946364]
27. Beacon 2.0. URL: https://beacon.anu.edu.au/ [accessed 2016-11-11] [WebCite Cache ID 6lvtqhYgD]
28. Mindhealthconnect. URL: https://www.mindhealthconnect.org.au/ [accessed 2016-11-11] [WebCite Cache ID 6lvtyQLVn]
29. Koohy H, Koohy B. A lesson from the ice bucket challenge: using social networks to publicize science. Front Genet 2014;5:430 [FREE Full text] [doi: 10.3389/fgene.2014.00430] [Medline: 25566317]
30. Barney LJ, Griffiths KM, Jorm AF, Christensen H. Stigma about depression and its impact on help-seeking intentions. Aust N Z J Psychiatry 2006 Jan;40(1):51-54. [doi: 10.1111/j.1440-1614.2006.01741.x] [Medline: 16403038]
31. Calear AL, Batterham PJ, Christensen H. Predictors of help-seeking for suicidal ideation in the community: risks and opportunities for public suicide prevention campaigns. Psychiatry Res 2014 Nov 30;219(3):525-530. [doi: 10.1016/j.psychres.2014.06.027] [Medline: 25048756]
32. Livingston JD, Boyd JE. Correlates and consequences of internalized stigma for people living with mental illness: a systematic review and meta-analysis. Soc Sci Med 2010 Dec;71(12):2150-2161. [doi: 10.1016/j.socscimed.2010.09.030] [Medline: 21051128]
33. Watson AC, Corrigan P, Larson JE, Sells M. Self-stigma in people with mental illness. Schizophr Bull 2007 Nov;33(6):1312-1318 [FREE Full text] [doi: 10.1093/schbul/sbl076] [Medline: 17255118]
34. Livingston JD, Cianfrone M, Korf-Uzan K, Coniglio C. Another time point, a different story: one year effects of a social media intervention on the attitudes of young people towards mental health issues. Soc Psychiatry Psychiatr Epidemiol 2014 Jun;49(6):985-990. [doi: 10.1007/s00127-013-0815-7] [Medline: 24401914]
35. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of Web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br J Psychiatry 2004 Oct;185:342-349 [FREE Full text] [doi: 10.1192/bjp.185.4.342] [Medline: 15458995]
36. Corrigan PW, Morris SB, Michaels PJ, Rafacz JD, Rüsch N. Challenging the public stigma of mental illness: a meta-analysis of outcome studies. Psychiatr Serv 2012 Oct;63(10):963-973. [doi: 10.1176/appi.ps.201100529] [Medline: 23032675]
37. Reiss D, Price RH. National research agenda for prevention research. The National Institute of Mental Health Report. Am Psychol 1996 Nov;51(11):1109-1115. [Medline: 8937258]
38. Leigh S. Comparing applets and oranges: barriers to evidence-based practice for app-based psychological interventions. Evid Based Ment Health 2016 Aug;19(3):90-92. [doi: 10.1136/eb-2016-102384] [Medline: 27431656]
39. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord 2015 Nov;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]
40. Inkster B, Stillwell D, Kosinski M, Jones P. A decade into Facebook: where is psychiatry in the digital age? Lancet Psychiatry 2016 Nov;3(11):1087-1090. [doi: 10.1016/S2215-0366(16)30041-4] [Medline: 27794373]
41. O'Dea B, Wan S, Batterham P, Calear A, Paris C, Christensen H. Detecting suicidality on Twitter. Internet Interv 2015;2(2):183-188.
42. Warden D, Rush AJ, Carmody TJ, Kashner TM, Biggs MM, Crismon ML, et al. Predictors of attrition during one year of depression treatment: a roadmap to personalized intervention. J Psychiatr Pract 2009 Mar;15(2):113-124 [FREE Full text] [doi: 10.1097/01.pra.0000348364.88676.83] [Medline: 19339845]

43. Gearing RE, Schwalbe CS, Short KD. Adolescent adherence to psychosocial treatment: mental health clinicians' perspectives on barriers and promoters. Psychother Res 2012;22(3):317-326. [doi: 10.1080/10503307.2011.653996] [Medline: 22313513]
44. Brown M, O'Neill N, van Woerden H, Eslambolchilar P, Jones M, John A. Gamification and adherence to Web-based mental health interventions: a systematic review. JMIR Ment Health 2016;3(3):e39 [FREE Full text] [doi: 10.2196/mental.5710] [Medline: 27558893]
45. Christensen H, Griffiths KM, Farrer L. Adherence in Internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
46. Aggarwal NK, Pieh MC, Dixon L, Guarnaccia P, Alegría M, Lewis-Fernández R. Clinician descriptions of communication strategies to improve treatment engagement by racial/ethnic minorities in mental health services: a systematic review. Patient Educ Couns 2016 Feb;99(2):198-209. [doi: 10.1016/j.pec.2015.09.002] [Medline: 26365436]
47. Oruche UM, Downs S, Holloway E, Draucker C, Aalsma M. Barriers and facilitators to treatment participation by adolescents in a community mental health clinic. J Psychiatr Ment Health Nurs 2014 Apr;21(3):241-248. [doi: 10.1111/jpm.12076] [Medline: 23682756]
48. Brown A, Rice SM, Rickwood DJ, Parker AG. Systematic review of barriers and facilitators to accessing and engaging with mental health care among at-risk young people. Asia Pac Psychiatry 2016 Mar;8(1):3-22. [doi: 10.1111/appy.12199] [Medline: 26238088]
49. French R, Reardon M, Smith P. Engaging with a mental health service: perspectives of at-risk youth. Child and Adolescent Soc Work J 2003 Dec;20(6):529-548. [doi: 10.1023/B:CASW.0000003142.13457.0a]
50. Gee A, McGarty C, Banfield M. Barriers to genuine consumer and carer participation from the perspectives of Australian systemic mental health advocates. J Ment Health 2016 Jun;25(3):231-237. [doi: 10.3109/09638237.2015.1124383] [Medline: 26697957]
51. Woodall A, Morgan C, Sloan C, Howard L. Barriers to participation in mental health research: are there specific gender, ethnicity and age related barriers? BMC Psychiatry 2010;10:103 [FREE Full text] [doi: 10.1186/1471-244X-10-103] [Medline: 21126334]
52. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. Aust N Z J Psychiatry 2009 Jun;43(6):571-578. [doi: 10.1080/00048670902873722] [Medline: 19440890]
53. Adair CE, McDougall GM, Beckie A, Joyce A, Mitton C, Wild CT, et al. History and measurement of continuity of care in mental health services and evidence of its role in outcomes. Psychiatr Serv 2003 Oct;54(10):1351-1356. [doi: 10.1176/appi.ps.54.10.1351] [Medline: 14557520]
54. Burns J, Hickie IB, Christensen EM. Strategies for adopting and strengthening e-mental health. URL: http:/ /nswmentalhealthcommission.com.au/sites/default/files/assets/File/ Report%20-%20The%20Sax%20Institute%20E-Mental%20Health%20Evidence%20Review%20cover%20page.pdf [accessed 2017-02-14] [WebCite Cache ID 6mAkM57qn]
55. Orlowski S, Lawn S, Matthews B, Venning A, Wyld K, Jones G, et al. The promise and the reality: a mental health workforce perspective on technology-enhanced youth mental health service delivery. BMC Health Serv Res 2016 Oct 10;16(1):562 [FREE Full text] [doi: 10.1186/s12913-016-1790-y] [Medline: 27724951]
56. Kushniruk A, Nøhr C. Participatory design, user involvement and health IT evaluation. Stud Health Technol Inform 2016;222:139-151. [Medline: 27198099]
57. 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;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
58. Kipping S, Stuckey MI, Hernandez A, Nguyen T, Riahi S. A Web-based patient portal for mental health care: benefits evaluation. J Med Internet Res 2016 Nov 16;18(11):e294 [FREE Full text] [doi: 10.2196/jmir.6483] [Medline: 27852556]
59. 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]
60. E-mental health strategy for Australia.: Australian Government Department of Health and Ageing; 2012. URL: http://www.health.gov.au/internet/main/publishing.nsf/content/7C7B0BFEB985D0EBCA257BF0001BB0A6/$File/emstrat.pdf [accessed 2017-02-14] [WebCite Cache ID 6lpoBwlSX]
61. Mobile medical applications: guidance for industry and food and drug administration staff.: U.S. Department of Health and Human Services: Food and Drug Administration URL: http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf [accessed 2016-11-21] [WebCite Cache ID 6mAreLcJo]
62. Larsen ME, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. PLoS One 2016;11(4):e0152285 [FREE Full text] [doi: 10.1371/journal.pone.0152285] [Medline: 27073900]
63. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]
64. Wicks P, Chiauzzi E. “Trust but verify”—five approaches to ensure safe medical apps. BMC Med 2015;13:205 [FREE Full text] [doi: 10.1186/s12916-015-0451-z] [Medline: 26404791]
65. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2014 Aug 21:e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]

66. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
67. Kelley PG, Bresee J, Cranor LF, Reeder RW. A nutrition label for privacy. 2009 Presented at: 5th Symposium on Usable Privacy and Security, SOUPS; July 15-17; Mountain View, California, USA.
68. Neff G. Why big data won't cure us. Big Data 2013 Sep;1(3):117-123 [FREE Full text] [doi: 10.1089/big.2013.0029] [Medline: 25161827]
69. Larson EB. Building trust in the power of “big data” research to serve the public good. JAMA 2013 Jun 19;309(23):2443-2444. [doi: 10.1001/jama.2013.5914] [Medline: 23780455]
70. Lupton D. How does digital health feel? Towards research on the affective atmospheres of digital health. Digital Health 2016:1 (forthcoming) [FREE Full text]
71. Kennedy H. Post, Mine, Repeat: Social Media Data Mining Becomes Ordinary. London: Palgrave Macmillan; 2016.
72. Wyatt JC, Sullivan F. eHealth and the future: promise or peril? BMJ 2005 Dec 10;331(7529):1391-1393 [FREE Full text] [doi: 10.1136/bmj.331.7529.1391] [Medline: 16339252]
73. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
74. Cane J, Richardson M, Johnston M, Ladha R, Michie S. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. Br J Health Psychol 2015 Feb;20(1):130-150. [doi: 10.1111/bjhp.12102] [Medline: 24815766]
75. Clough BA, Casey LM. Smart designs for smart technologies: Research challenges and emerging solutions for scientist-practitioners within e-mental health. Prof Psychol—Res Pract 2015;46(6):429-436. [doi: 10.1037/pro0000053]
76. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34 Suppl:1220-1228. [doi: 10.1037/hea0000305] [Medline: 26651463]
77. Thornicroft G, Rose D, Huxley P, Dale G, Wykes T. What are the research priorities of mental health service users? J Ment Health 2009 Jul 06;11(1):1-3. [doi: 10.1080/096382301200041416]
78. Proudfoot J, Parker G, Hadzi PD, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]

Abbreviations

eMHPrac: eMental Health in Practice
GP: general practitioner
NHS: National Health Service
UNSW: University of New South Wales

Edited by K Dinakar, R Calvo; submitted 21.11.16; peer-reviewed by E Ludman, J Torous, G Wadley; comments to author 14.01.17; revised version received 29.01.17; accepted 04.02.17; published 28.02.17.

Please cite as:
Nicholas J, Huckvale K, Larsen ME, Basu A, Batterham PJ, Shaw F, Sendi S
Issues for eHealth in Psychiatry: Results of an Expert Survey
J Med Internet Res 2017;19(2):e55
URL: http://www.jmir.org/2017/2/e55/
doi:10.2196/jmir.6957
PMID:28246068

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org