Journal of Medical Internet Research

Contents

Viewpoint

Reviews

Letters to the Editor

Corrigenda and Addenda

Editorial

On Supplementing "Foot in the Door" Incentives for eHealth Program Engagement

Marc Steven Mitchell[1], MSc; Guy E Faulkner[1], PhD

Faculty of Kinesiology and Physical Education, Graduate Department of Exercise Sciences, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Marc Steven Mitchell, MSc
Faculty of Kinesiology and Physical Education
Graduate Department of Exercise Sciences
University of Toronto
Athletic Center
55 Harbord St
Toronto, ON, M5S 2W6
Canada
Phone: 1 416 660 7881
Fax: 1 416 597 7160
Email: marc.mitchell@uhn.ca

Related Article:

Comment on: http://www.jmir.org/2014/7/e163/

Abstract

Financial health incentives, such as paying people to lose weight, are being widely implemented by Western nations and large corporations. A growing number of studies have tested the impact of incentives on health behaviors, though few have evaluated the approach on a population-scale. In this issue of the Journal of Medical Internet Research, Liu et al add to the evidence-base by examining whether a single incentive can motivate enrollment and engagement in a preventive eHealth program in a sample of 142,726 Canadian adults. While the incentives increased enrollment significantly (by a factor of about 28), a very high level of program attrition was noted (90%). The "foot in the door" incentive technique employed was insufficient; enrollees received incentives for signing-up for, but not for engaging with, the eHealth program. To supplement this technique and drive sustained behavior change, several theoretically- and empirically-based strategies are proposed. Specifically, incentives indexed to behavioral achievements over time are highlighted as one approach to boost engagement in this population in the future.

Attrition (non-use or dropout) is one of the hallmarks of eHealth interventions [1] and a particular problem for online lifestyle interventions [2], but also in the offline setting. Financial health incentives, such as paying people to attend nutrition classes, get vaccinated, or lose weight, are being widely implemented by Western nations and large corporations [3-5]. A growing number of studies have tested the impact of incentives on "single-shot" (eg, clinic visits) [6] and "lifestyle" (eg, exercise) [7] health behaviors with promising results, though few have evaluated the approach on a population-scale. Notably, Stock et al conducted a unique evaluation of a large-scale incentive program in Germany (N=70,429) and found that modest incentives motivated participation in disease prevention programs, saving the German Statutory Health Insurance about

€100 per person per year in direct health costs [8]. In this issue of the *Journal of Medical Internet Research*, Liu et al add to the evidence-base by examining whether a single incentive exposure (20 loyalty points worth about Can $2) can motivate enrollment *and continued participation* in a preventive eHealth program in a sample of 142,726 Canadian adults [9].

Not surprisingly, and consistent with the "single-shot" incentives literature, the authors found that individuals offered the incentive were 28 times more likely to enroll in the eHealth program. The question is, was this impressive boost in enrollment meaningful considering the significant attrition (about 90%) noted by the authors? The data would suggest that this was the case. Although program attrition was substantial, the fact remains that about

5000 adults who might not have otherwise enrolled in the eHealth program (without the contingent incentive) continued onto the next step of the program (ie, assessing their readiness to change), and about 1000 of those individuals remained engaged 6 weeks later (ie, with a second round of assessing readiness to change). It is not clear whether this level of engagement (assessing readiness to change on two occasions and, presumably, reading tailored-education materials) was sufficient to stimulate "lifestyle" health behaviors to the extent required to produce clinically or health-economically relevant changes. This is a topic for future study (linking participation to administrative databases, for example). Regardless, the incentive approach employed here holds great promise for enhancing the delivery (reach) of Web-based prevention programs. In the midst of this potential, however, there is room to improve upon the design, and maximize the impact, of the incentive offering.

First, the "foot in the door" design technique employed was insufficient to drive eHealth program engagement in this population. While getting people to agree to do something small (put their "foot in the door"; ie, enroll in the eHealth program), can sometimes make them more likely to do something big (ongoing eHealth program use), according to behavioral economics [10], this is not what transpired here. Rather, it appears that "rewards simply motivated people to get rewards", an often-cited risk of incentives for health [11]. Once incentives are removed, it is said that people tend to revert to baseline behaviors or worse, if autonomous motivation is undermined by extrinsic rewards. New evidence suggests, however, that well-designed incentives can actually promote quality behavior change (ie, by protecting/building autonomous motivation), and increase the potential for sustained effects [12].

The authors recommend supplemental strategies to maintain engagement that stem from previous work in the area of incentives for exercise [7]. From the list of design suggestions mentioned, the most critical suggestion likely has to do with indexing incentives to behavioral occurrences over time. "Lifestyle" health behaviors (like exercise or eHealth program use) can be hard to adopt, in part, because the costs (eg, time) of engaging in these behaviors are usually borne in the present (carrying disproportionate weight in decision making, a phenomenon referred to as the "present bias") [10], and the benefits (eg, health) are often delayed and thus discounted. As a result, people tend to act in favor of their immediate self-interest, at the expense of their long-term wellbeing. Offering incentives for *regular* (perhaps daily or weekly) eHealth program use (eg, completing stage of change

assessments, or diet/exercise diaries) may increase the immediately rewarding aspects of participation and in turn people's propensity to log into and use the eHealth tool.

Regular incentive offerings need not be prohibitively costly either. In fact, with a stronger application of behavioral economic principles [13], and considering the full range of incentive design options [14], it may be possible to produce greater effects with the same investment ($77,000, or about $2 per enrollee). For example, it is worth exploring whether the proportion of Canadians engaged at 6 weeks would increase if 10 Air Miles (vs 20 Air Miles) were offered for enrollment, with the remaining Air Miles (ie, 10) offered for behavioral achievements over subsequent weeks (eg, 2 Air Miles per week fruit/vegetable intake self-reported). To further optimize the approach, a lottery could be layered on top of this assured incentive scheme, where participants who perform desired behaviors are entered into weekly draws to win additional Air Miles. This lottery feature could be used to exploit peoples' tendency to overweight small probabilities [10] and drive target behaviors without significant additional resource. Notably, according to Klein & Karlawish, older adults may be more sensitive to lottery-based incentives than their younger counterparts [15]. Also, tailoring incentive offerings to demographic groups may boost enrollment and engagement. For example, making "meaningful" incentives more salient, by marketing iTunes credits to teenagers, wellness holiday credits to working adults, or grocery store vouchers to seniors, may increase the impact of relatively modest rewards. The newly developed Health Incentive Program Questionnaire may be useful in identifying preferred, more meaningful, and valuable voucher options (Mitchell et al, under review).

Though the Liu et al [9] paper is constrained by several limitations (eg, selection bias, sub-optimal incentive design, self-reported outcomes) it is an important launching pad for more sophisticated population-level evaluations of incentives for health. Incentives have emerged as a practical and acceptable public health policy alternative in Canada (and elsewhere), and so to promote broad, effective, and sustained implementation, future evaluations should stretch beyond the question, "Do incentives work?" to questions of moderations, such as, "Who or what behaviors are most sensitive to incentive intervention?", "What types or sizes of incentives work best or are necesary?", and "How long must incentives be in place to produce long-lasting change?". Learning more about how to design incentives that produce clinically or health-economically significant changes, without being prohibitively costly, will only help to optimize this promising intervention in the future.

References

1. 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]
2. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
3. Ries NM. Financial incentives for weight loss and healthy behaviours. Healthc Policy 2012 Feb;7(3):23-28 [FREE Full text] [Medline: 23372578]

4.   National Business Group on Health. 2010 Aug. Large Employers&#8217; 2011 Health Plan Design Changes URL: http://www.businessgrouphealth.org/pub/f314abb8-2354-d714-51f7-315a5a14fa20 [WebCite Cache ID 6RKJYOhim]

5.   Madison K, Schmidt H, Volpp KG. Smoking, obesity, health insurance, and health incentives in the Affordable Care Act. JAMA 2013 Jul 10;310(2):143-144. [doi: 10.1001/jama.2013.7617] [Medline: 23765171]

6.   Jochelson K. Kicking Bad Habits. 2007. Paying the patient: Improving health using financial incentives URL: http://www.kingsfund.org.uk/sites/files/kf/field/field_document/paying-the-patient-kicking-bad-habits-supporting-paper-karen-jochelson.pdf [WebCite Cache ID 6RKJcd0lP]

7.   Mitchell MS, Goodman JM, Alter DA, John LK, Oh PI, Pakosh MT, et al. Financial incentives for exercise adherence in adults: systematic review and meta-analysis. Am J Prev Med 2013 Nov;45(5):658-667. [doi: 10.1016/j.amepre.2013.06.017] [Medline: 24139781]

8.   Stock S, Schmidt H, Büscher G, Gerber A, Drabik A, Graf C, et al. Financial incentives in the German Statutory Health Insurance: new findings, new questions. Health Policy 2010 Jun;96(1):51-56. [doi: 10.1016/j.healthpol.2009.12.015] [Medline: 20106543]

9.   Liu S, Hodgson C, Zbib AM, Payne AY, Nolan RP. The effectiveness of loyalty rewards to promote the use of an internet-based heart health program. J Med Internet Res 2014;16(7):e163 [FREE Full text] [doi: 10.2196/jmir.3458] [Medline: 24989982]

10. Camerer C, Loewenstein G, Rabin M. Behavioral Economics: Past, Present, Future. In: Advances in Behavioral Economics. Princeton: Princeton University Press; 2003.

11. Moller AC, McFadden HG, Hedeker D, Spring B. Financial motivation undermines maintenance in an intensive diet and activity intervention. J Obes 2012;2012:740519 [FREE Full text] [doi: 10.1155/2012/740519] [Medline: 22548152]

12. Promberger M, Marteau TM. When do financial incentives reduce intrinsic motivation? comparing behaviors studied in psychological and economic literatures. Health Psychol 2013 Sep;32(9):950-957 [FREE Full text] [doi: 10.1037/a0032727] [Medline: 24001245]

13. Loewenstein G, Asch DA, Volpp KG. Behavioral economics holds potential to deliver better results for patients, insurers, and employers. Health Aff (Millwood) 2013 Jul;32(7):1244-1250. [doi: 10.1377/hlthaff.2012.1163] [Medline: 23836740]

14. Adams J, Giles EL, McColl E, Sniehotta FF. Carrots, sticks and health behaviours: a framework for documenting the complexity of financial incentive interventions to change health behaviours. Health Psychol Rev 2014 Sep;8(3):286-295. [doi: 10.1080/17437199.2013.848410] [Medline: 25053215]

15. Klein E, Karlawish J. Challenges and opportunities for developing and implementing incentives to improve health-related behaviors in older adults. J Am Geriatr Soc 2010 Sep;58(9):1758-1763 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03030.x] [Medline: 20863335]

XSL•FO
RenderX

The Effectiveness of Loyalty Rewards to Promote the Use of an Internet-Based Heart Health Program

Sam Liu[1,2], MSc; Corinne Hodgson[3], MSc; Ahmad M Zbib[4], MD; Ada YM Payne[1], PhD; Robert P Nolan[1,2], PhD

[1]Cardiac eHealth Research Unit, Peter Munk Cardiac Center, University Health Network, Toronto, ON, Canada

[2]Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[3]Corinne S Hodgson & Associates, Burlington, ON, Canada

[4]Digital Health and Innovation, Heart and Stroke Foundation of Canada, Toronto, ON, Canada

Corresponding Author:
Robert P Nolan, PhD
Cardiac eHealth Research Unit
Peter Munk Cardiac Center
University Health Network
Room 6N-618 NU
585 University Avenue
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4800 ext 6400
Fax: 1 416 340 3162
Email: rnolan@uhnres.utoronto.ca

Related Article:

http://www.jmir.org/2014/7/e179/

Abstract

Background: Internet-based health programs have been shown to be effective in reducing risk for cardiovascular disease. However, their rates of enrollment and engagement remain low. It is currently unclear whether rewards from established loyalty programs can serve as a conditioned stimulus to improve the use of a freely available Internet-based program.

Objective: The objectives of the study were to (1) examine enrollment rates and levels of engagement with the My Health eSupport program between a Conditioned Reward group and a Control group, and (2) investigate the influence of loyalty rewards and participant characteristics on levels of enrollment and program engagement.

Methods: The study sample (n=142,726) consisted of individuals who were offered enrollment in an Internet-based health intervention (My Health eSupport) after completing the Heart&Stroke Risk Assessment on the Heart and Stroke Foundation website. My Health eSupport programs provided encouragement and tips for lifestyle change. This is a free, self-guided, fully automated program that proactively delivers tailored email messages at 2-week intervals based on the participant's stage of motivational "readiness" and priority for lifestyle change. Participants in the Conditioned Reward group were offered a single exposure of 20 loyalty reward points from the Air Miles loyalty program for completing the Heart&Stroke Risk Assessment (10 reward points) and enrolling in the Internet-based program (10 reward points). Meanwhile, no rewards were given to the Control group participants. All data were collected between February 1, 2011 and February 10, 2012.

Results: In total, 51.38% (73,327/142,726) of individuals in the Conditioned Reward group and 48.62% (69,399/142,726) of individuals in the Control group completed the Heart&Stroke Risk Assessment. Subsequently, significantly more individuals from the Conditioned Reward group (52.96%, 38,835/73,327) enrolled in the My Health eSupport program than Controls (4.07%, 2826/69,399). Regression analyses indicated that individuals were 27.9 times (95% CI 26.4-29.4; $P<.001$) more likely to join the My Health eSupport program when presented with loyalty rewards controlling for gender, age, education, ethnicity, employment, and number of modifiable risk factors. However, ongoing engagement level was low in both groups and it was not influenced by loyalty rewards. Instead, individuals were more likely to engage with the My Health eSupport program if they were greater than 60 years of age (OR 12.56, 95% CI 5.66-27.8; $P<.001$), were female (OR 1.27, 95% CI 1.09-1.46; $P=.002$), or had one or more modifiable risk factors (OR 1.38, 95% CI 1.31-1.45; $P<.001$).

Conclusions: Our findings suggest that a single exposure of loyalty rewards may be used to encourage individuals to enroll in an Internet-based preventative health program, but additional strategies are required to maintain engagement level. Future studies need to examine the schedules of loyalty reward reinforcement on the long-term engagement level of Internet-based health programs.

Introduction

Cardiovascular diseases (CVDs) are the number one cause of death globally. It is estimated that the number of people who die from CVDs will reach 23.3 million by 2030 [1]. Studies have shown that more than 80% of CVDs can be prevented through lifestyle changes, such as regular exercise and a healthy diet [2].

The wide adoption of Internet usage presents an incredible opportunity for delivering preventive health initiatives at a population level. In 2012, approximately 83% of Canadians had personal access to the Internet: 85% of whom live in metropolitan areas and 75% in rural areas [3]. This includes 63-78% of Canadians who are in the lowest income quartiles and 84% of Canadians between 45 and 64 years, which is an age group with increased risk for CVD. Several randomized controlled trials and meta-analyses have shown that Internet-based lifestyle interventions can help individuals improve self-care behaviors (physical activity, dietary habits) [4,5], psychological functioning (anxiety, quality of life) [6], and clinical outcomes (blood pressure, weight, blood glucose) [7-9]. However, enrollment and engagement in Internet-based health interventions remain a challenge [10-12]. This can limit the significant impact that Internet-based interventions have on health behavior change at a population level.

External rewards such as financial and loyalty rewards have been used as a strategy to increase the enrollment and engagement of behavioral interventions [13,14]. The theoretical rationale for using these external rewards to reinforce behavior change is based on both operant learning and behavioral economic theory [13,15]. The key principle of operant learning is that behavior is conditioned by contingencies that are rewarding or punishing in nature [13]. Economic theory includes the principles from operant theory and states that an individual tends to behave in ways that maximize one's immediate reward [13,15]. According to these theories, the propensity for enrollment and sustained engagement in Internet-based programs may be enhanced by introducing external rewards. On the contrary, self-determination theory (SDT) asserts that an external reward may weaken intrinsic motivation for behavior change and this is known as the "crowding out" effect [16,17]. It is currently unclear whether a single bout of reward from a loyalty program can serve as a conditioned stimulus affecting the enrollment and engagement of a freely available Internet-based health program. An advantage of using loyalty rewards over financial ones is that it may be more sustainable as both businesses and consumers can mutually benefit. Loyalty reward programs can help businesses gain new and retain existing customers by rewarding them with loyalty "reward points". These customers can then redeem their reward points for goods and/or services from the loyalty program. Thus, the aims of this study were to (1) examine enrollment rates and levels of engagement with the My Health eSupport program between the Conditioned Reward and the Control group, and (2) investigate the influence of loyalty rewards and participant characteristics on levels of enrollment and program engagement. A previous study reported that a single bout of financial reward (US$2.00-$20.00) could increase enrollment and response to a follow-up survey 3 months after enrollment for an Internet-based health program [18]. Therefore, we hypothesized that enrollment rates and level of engagement would be higher in the Conditioned Reward group than Controls. Results were expected to be influenced by age, ethnicity, education, gender, employment status, and the number of modifiable risk factors [11,18].

Methods

Overview

This was an observational study and the data was collected by the Heart and Stroke Foundation of Canada (HSFC) between February 1, 2011 and February 10, 2012. The HSFC is a non-profit public organization that aims to empower Canadians to live healthy lives free of heart disease and stroke by raising awareness of the key risk factors and encouraging and supporting them to play an active role in managing their health. Study participants were 18 years or older and provided consent for the study. All personal identifiers were removed prior to retrieving the records from the HSFC database. This study was approved by the University Health Network Ethics Board.

Recruitment and Study Protocol

The study sample consisted of individuals who were invited to participate in an Internet-based heart health program (My Health eSupport) after completing their Heart&Stroke Risk Assessment on the HSFC website. A total of 142,726 individuals completed the Heart&Stroke Risk Assessment and were included in our analyses. This sample was comprised of two groups: Conditioned Reward vs Control. Participants in the Conditioned Reward group were recruited from the Air Miles loyalty program using a standardized recruitment email. The email described the opportunity to receive a non-cash, uniform reward of up to 20 Air Miles reward points for enrolling in the My Health eSupport HSFC program. Specifically, participants were assured of immediately receiving 10 Air Miles reward points for completing the Heart&Stroke Risk Assessment and another 10 Air Miles for enrolling in My Health eSupport. No additional

rewards were given after enrollment. The Heart&Stroke Risk Assessment was a free e-tool on the HSFC website that first enabled individuals to assess their lifestyle in regard to CVD risk. It then offered enrollment to the My Health eSupport program. Enrollment in the My Health eSupport program required the participants to first complete the Heart&Stroke Risk Assessment and create a log-in ID on the HSFC website. Air Miles participants accessed the Heart&Stroke Risk Assessment and claimed the reward using a unique Web link embedded in the recruitment email. Air Miles participants who completed the Heart&Stroke Risk Assessment were assigned to the Conditioned Reward group. Participants in the Control group accessed the Heart&Stroke Risk Assessment on the HSFC website without using the unique Web link provided by the Air Miles loyalty program. Control participants did not receive any rewards for completing either program.

Program Description

The My Health eSupport program was a free, self-guided, fully automated healthy lifestyle program that proactively delivered tailored email messages at 2-week intervals. The emails contained information on heart healthy living with links to the HSFC website. The initial email from My Health eSupport was delivered to the participants following enrollment. The email guided participants to report their stage of motivational "readiness" to adhere to Health Canada guidelines for diet (daily intake of fruit, vegetables, and restriction of dietary fat and salt), exercise (planned exercise and daily activity), and smoke-free living. Readiness for change was operationally defined for each behavior according to the conventional algorithm of Prochaska's Transtheoretical Model [19] (Pre-contemplation, Contemplation, Preparation, Action, and Maintenance). In an effort to reinforce motivation and efficacy for self-directed change [20], subjects were then prompted to choose their priority for lifestyle change from the above-noted behaviors associated with diet, exercise, and smoke-free living. Subsequently, three email messages tailored to the individuals' readiness and priority for lifestyle change were delivered at 2-week intervals. The assessment of readiness and priority for lifestyle change was then repeated 6 weeks later from the initial assessment.

Measures

This paper evaluated enrollment rates and levels of engagement in the My Health eSupport program in both the Conditioned Reward and Control groups. Enrollment rate was calculated as the number of individuals who enrolled in the My Health eSupport program divided by the number of participants who completed the Heart&Stroke Risk Assessment. Two measures of engagement level with the program were calculated. The initial engagement was defined as the proportion of individuals who completed the assessment of readiness and priority for lifestyle change during the first email. Ongoing engagement was defined as completion of the assessment of readiness and priority for lifestyle the second time at 6 weeks following initial engagement. Participants completed the second assessment of readiness and priority for lifestyle change within 8 weeks from

receiving the re-assessment email. Characteristics of participants were extracted from self-reported data in the Heart&Stroke Risk Assessment, which is an online self-assessment questionnaire. Characteristics included age, gender, education level, ethnicity, employment, medical condition(s), and body mass index (calculated from height and weight). Modifiable CVD risk factors were defined as the following: physical activity level (whether participants achieved 30-60 minutes of moderate exercise, 4 times per week), smoking status (Yes or No), excess alcohol consumption (Male: consumed >2 drinks a day or >14 drinks a week; Female: consumed >2 drinks a day, >9 drinks a week), and whether participants consumed high fat foods (3 or more times per week), fruits and vegetable (5 or more servings per day), and food with high salt content (3 or more times per week).

Data Analysis

Chi-square and independent t tests were used to compare characteristic differences between the Conditioned Reward and Control groups. Statistical significance can be influenced by a large sample size [21]; therefore, we calculated effect size using Cramer's V (post chi-square statistic) and Cohen's d (post independent t test) to determine the strength of these statistically significant results. A weak, moderate, and strong association (effect size) for Cramer's V was defined as <0.2, 0.2-0.6, and >0.6, respectively [22]. Meanwhile, a small, moderate, and large effect size for Cohen's d was defined as ≤0.2, 0.5, and ≥0.8, respectively [23]. Binary logistic regressions were conducted to estimate the association between loyalty rewards and enrollment (model 1), initial engagement (model 2), and ongoing engagement (model 3) of the Internet-based health program. Odds ratios were calculated to identify the effects of loyalty rewards on the likelihood of enrolling and engaging with the My Health eSupport program. Based on previous research [11,18], all models were adjusted for age, ethnicity, education, gender, employment status, and the number of modifiable risk factors. Data were analyzed using SPSS version 19 and significance was accepted at an alpha level of .01.

Results

Participants

A total of 73,327 individuals in the Conditioned Reward group and 69,399 individuals in the Control group completed the Heart&Stroke Risk Assessment. Baseline characteristics of the participants are presented in Table 1. Overall, both groups attracted a larger proportion of females than males (68.48%, 97,732/142,726 vs. 31.52%, 44,994/142,726). Most of the participants were Caucasian (83.54%, 119,239/142,726). The most common illnesses in both groups were hypertension (22.43%, 32,014/142,726) and dyslipidemia (17.86%, 25,487/142,726). More than 50% of individuals reported having a BMI higher than 25. Two of the most common lifestyle factors associated with CVD risk for both groups were physical inactivity (44.10%, 62,942/142,726) and inadequate fruit and vegetable consumption (43.32%, 61,828/142,726) (Table 2).

Table 1. Baseline participant characteristics.

Characteristics		Conditioned Rewardn=73,327 n (%)	Controln=69,399 n (%)	*P* value	Effect size
Age (years), mean (SD)		50.4 (14.2)	50.5 (14.4)	.54	0.003
Female		49,778 (67.88%)	47,954 (69.11%)	<.001	0.01
Completed university education		43,783 (60.74%)	41,208 (60.53%)	.40	0.002
Ethnicity				<.001	0.13
	Caucasian	62,658 (85.48%)	56,581 (81.62%)		
	South Asian	1688 (2.30%)	1911 (2.76%)		
	Chinese	3040 (4.15%)	1236 (1.78%)		
	Aboriginal	1038 (1.42%)	1253 (1.81%)		
	African/black	683 (0.93%)	1160 (1.67%)		
Marital status				<.001	0.03
	Married	42,396 (57.85%)	40365 (58.18%)		
	Widowed, divorced	10,177 (13.89%)	8423 (12.14%)		
	Single	10,787 (14.72%)	10,397 (14.99%)		
Employment					
	Full or part-time	41,279 (56.33%)	40,345 (58.16%)	<.001	0.07
	Self-employed	6675 (9.11%)	6726 (9.70%)		
	Retired	14,644 (19.98%)	11,654 (16.80%)		
	Stay-at-home parent	3049 (4.16%)	2350 (3.39%)		
Type of employment				<.001	0.08
	Management/white collar	45,660 (63.00%)	43,364 (68.81%)		
	Sales or services	9491 (13.09%)	6757 (10.72%)		
	Trades	6408 (8.84%)	6082 (9.65%)		
Body Mass Index				<.001	0.13
	Underweight (<18.5)	1464 (2.00%)	1109 (1.60%)		
	Normal weight (18.5-24.9)	21,942 (29.93%)	19,384 (27.94%)		
	Overweight (25-29.9)	21,240 (28.97%)	22,356 (32.21%)		
	Obesity (≥30)	16,459 (22.45%)	20,039 (28.88%)		
Mean number of cardiovascular co-morbidities (SD)		0.54 (0.93)	0.51 (0.89)	<.001	0.03
	Diabetes	5298 (7.23%)	3072 (4.43%)	<.001	0.06
	Dyslipidemia	13,984 (19.07%)	11,503 (16.58%)	<.001	0.03
	Heart disease	2795 (3.81%)	2662 (3.84%)	.81	0.001
	Hypertension	15,520 (21.17%)	16,494 (23.77%)	<.001	0.11
	Stroke	1254 (1.71%)	1327 (1.91%)	.004	0.01
	Mood disorder	11,906 (16.24%)	8857 (12.76%)	<.001	0.05
	Sleep apnea	3988 (5.44%)	3065 (4.42%)	<.001	0.02

Table 2. Baseline modifiable risk factors.

Modifiable risk factors	Conditioned Reward n=73,327 n (%)	Control n=69,399 n (%)	P value	Effect size
Physical inactivity	30,972 (42.24%)	31,970 (46.07%)	<.001	0.04
Smoker	5704 (8.22%)	8270 (11.30%)	<.001	0.05
Excess alcohol	17,095 (23.31%)	16,615 (23.94%)	.005	0.01
Fatty foods	8537 (11.69%)	10,663 (15.42%)	<.001	0.06
Infrequent fruit and vegetable	30,601 (41.82%)	31,227 (45.05%)	<.001	0.03
Salt	19,738 (26.92%)	14,205 (20.47%)	<.001	0.08
Mean number of modifiable risk factors (SD)	2.51 (1.37)	2.58 (1.37)	<.001	0.05

Enrollment Rates

Enrollment in the My Health eSupport program was higher in the Conditioned Reward group (52.96%, 38,835/73,327) than the Controls (4.07%, 2826/69,399). Loyalty rewards provision was the strongest predictor for enrollment in the My Health eSupport program (OR 27.9, 95% CI 26.4-29.4; P<.001). Factors influencing enrollment are presented in Table 3.

Table 3. Prediction of enrollment and engagement.

Factors	Enrollment			Initial engagement			Ongoing engagement		
	OR	(95% CI)	P value	OR	(95% CI)	P value	OR	(95% CI)	P value
Loyalty rewards (Conditioned Reward vs Control)	27.9	(26.4-29.4)	<.001	0.79	(0.66-0.92)	.004	0.93	(0.68-1.26)	.662
Gender (female vs male)	0.96	(0.93-0.99)	.007	1.26	(1.17-1.36)	<.001	1.27	(1.09-1.46)	.002
Age									
≤29 years	1.98	(1.82-2.15)	<.001	1.37	(0.95-1.97)	.085	1.49	(0.62-3.55)	.371
30-39 years	1.69	(1.56-1.82)	<.001	2.32	(1.64-3.27)	<.001	1.99	(0.86-4.57)	.104
40-49 years	1.80	(1.66-1.94)	<.001	4.41	(3.14-6.17)	<.001	5.86	(2.63-13.0)	<.001
50-59 years	1.80	(1.66-1.94)	<.001	6.81	(4.87-9.49)	<.001	9.45	(4.27-20.8)	<.001
≥60 years	2.12	(1.96-2.30)	<.001	8.40	(6.00-11.7)	<.001	12.56	(5.66-27.8)	<.001
Ethnicity (Caucasian vs other)	1.20	(1.15-1.25)	<.001	1.13	(1.01-1.24)	.025	1.12	(0.89-1.38)	.318
Total number of modifiable risk factors	0.96	(0.95-0.97)	<.001	1.37	(1.33-1.40)	<.001	1.38	(1.31-1.45)	<.001
Employment (employed vs not employed)	1.02	(0.99-1.06)	.201	0.89	(0.82-0.95)	<.001	1.03	(0.88-1.20)	.676
University education (completed vs not completed)	1.05	(1.01-1.08)	.004	0.86	(0.79-0.91)	<.001	0.88	(0.76-1.01)	.068

Engagement Level

After participants enrolled in the My Health eSupport program, only 12.43% (4829/38,835) of individuals in the Conditioned Reward group and 8.49% (240/2826) of the Controls assessed their readiness for change and selected a priority area for lifestyle change. Out of these participants, 20.98% (1013/4829) in the Conditioned Reward group and 24.17% (58/240) in the Control group completed the second assessment at week 6. In our regression analyses, loyalty rewards strategy was negatively associated with initial engagement (OR 0.79, 95% CI 0.66-0.92; P=.004) and it was not a significant predictor for ongoing engagement (P=.66). Meanwhile, age, gender, and numbers of modifiable risk factors were significant predictors of ongoing engagement (Table 3).

Discussion

Principal Findings

The main finding of this study was that a single exposure of loyalty rewards significantly influenced enrollment for the My Health eSupport program. Individuals were 27.9 times more likely to enroll when presented with loyalty rewards. However, contrary to our hypothesis, ongoing engagement was not influenced by loyalty rewards. Individuals were more likely to engage with the program if they were greater than 60 years of age, female, or had one or more modifiable risk factors. These findings suggest that a single exposure of loyalty rewards may be used to encourage individuals to sign up for an Internet-based preventative health initiative, but supplemental strategies are required to maintain engagement.

Our findings on the effects of loyalty rewards with online enrollment were consistent with other study findings using

XSL•FO
RenderX

financial rewards [24]. Alexander et al [18] reported that financial incentives of $2.00, $5.00, and $10.00 offered at the point of enrollment were able to enroll 3.5%, 7.5%, and 2.7% of the participants, respectively. It is encouraging to observe that, in this study, 20 Air Miles (equivalent to about $2.00 CAD) was sufficient to enroll approximately 50% of the participants. It is important to take into consideration that the participants in the Conditioned Reward group were from the Air Miles loyalty reward program. These individuals may have perceived the value of the loyalty reward as being greater than the direct monetary value of $2.00. As a result, our findings may be limited in the ability to generalize beyond Air Miles members.

Engagement level with Internet-based interventions has always been a challenge as high dropout and losses to follow-up are common. This phenomenon is described by Eysenbach as "the law of attrition" [25]. It is hypothesized that the attrition curve can be a valuable marker of the underlying causes for attrition. In our study, the attrition followed an L-shaped curve, which reflected an initial rapid decline of participants, with a remaining group of steady "hardcore" users who continuously used the system over the 6-week interval of this study. This type of attrition curve may indicate that the enrolled participants may not be the appropriate user group for the HSFC website [25] and/or that participants did not perceive the benefits of continued engagement with the Internet-based health intervention [26]. Since the My Health eSupport program was designed to reduce CVD risk through lifestyle changes, it may not be surprising that the minority of individuals who continued to engage with the eSupport program were those with higher CVD risk, such as older individuals or individuals with one or more modifiable CVD risk factors.

A single exposure of loyalty reward may have also contributed to the low level of engagement with the My Health eSupport program. Since rewards were presented at the point of enrollment, some individuals in the Conditioned Reward group may have enrolled for the sole purpose of obtaining this reward. This may explain the negative association found between loyalty rewards and initial engagement level. Charness et al [27] found that financial rewards may need to be presented on a number of occurrences (for at least 5 weeks) in order to help individuals pass the "threshold" level needed to sustain an activity on their own for up to 16 weeks. The long-term effectiveness (>6 months) of using financial incentives to maintain behavior change remains unclear, primarily due to the fact that the majority of studies are short term (<6 months) [17]. A recent meta-analysis reported several design features for financial rewards that may help increase adherence to exercise programs. These included (1) providing an assured financial incentive rather than lottery or chance-based reward, (2) offering larger incentives, and (3) incorporating incentives that are designed to be indexed (eg, $25/pound of weight loss), or escalating (eg, $10 for the first pound lost, $15 for second pound lost, etc) [17]. It is important to incorporate these design features into future studies aimed at behavior change using loyalty rewards.

Some researchers have cautioned against the use of any external incentives as they may have the undesired effects of inhibiting intrinsic motivation [27,28]. Future studies on Internet-based interventions should consider how indexed or escalated loyalty rewards could be used to sustain program engagement while maintaining or increasing intrinsic motivation associated with a health behavior change. Based on self-determination theory, intrinsic motivation is fulfilled by the psychological needs of competence (a sense of mastery), autonomy (ownership over behavior), and social relatedness (feel socially connected to others) [29]. Mitchell et al suggested that extrinsic rewards may be used to fulfill these psychological needs in order to avoid harming intrinsic motivation by (1) rewarding achievements of realistic self-regulatory goals (eg, the use of self-monitoring), (2) providing choice for the types of reward and the activities chosen, and (3) offering rewards that are related to social outcomes (eg, group contingencies or charitable donations) [17]. Additionally, it may be necessary to combine loyalty rewards with previously established key components of counselling in order to maintain levels of engagement in Internet-based interventions. These components include: goal setting and self-monitoring; the provision of feedback based on the patients' activity; self-efficacy enhancement; and relapse prevention [30].

Limitations and Strengths

Several limitations of the present study should be noted. The accuracy of self-reported data is open to challenge for validity. Self-report bias such as lifestyle behaviors of exercise, diet, and smoking could have influenced the accuracy of our results. Engagement with the system was defined when participants re-assessed their readiness for change and selected another priority area for lifestyle change. It is possible that individuals maintained engagement with the program but never completed the re-assessment, which may underestimate the levels of engagement with the My Health eSupport Program. There are other variables that may be associated with levels of enrollment and engagement which were not assessed, such as income, anxiety, and depression. Finally, there may be selection bias in our study as the Conditioned Reward group consisted of participants from the Air Miles loyalty reward program. This may limit our ability to generalize our findings beyond Air Miles members. A strength of this study includes the large sample size. This is one of the first "real-world" (eg, population-based) studies that has been conducted to examine the effects of loyalty rewards on enrollment and engagement with an Internet-based intervention.

Conclusions

Internet-based interventions hold great potential for delivering preventive health initiatives at a population level. A single exposure of loyalty rewards can increase enrollment but additional strategies are required to maintain engagement level. This study has significant design implications for incorporating loyalty rewards as an effective enrollment strategy for future Internet-based interventions. More research is needed that explores the long-term effects of using loyalty rewards to reinforce engagement level and the associated efficacy of Internet-based health programs.

Conflicts of Interest

None declared.

References

1. 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]
2. Buttar HS, Li T, Ravi N. Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation. Exp Clin Cardiol 2005;10(4):229-249 [FREE Full text] [Medline: 19641674]
3. Statistics Canada. 2013. Individual Internet use and e-commerce URL: http://www.statcan.gc.ca/daily-quotidien/131028/dq131028a-eng.pdf [accessed 2014-04-06] [WebCite Cache ID 6OdipGQm7]
4. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
5. Yoo HJ, Park MS, Kim TN, Yang SJ, Cho GJ, Hwang TG, et al. A ubiquitous chronic disease care system using cellular phones and the internet. Diabet Med 2009 Jun;26(6):628-635. [doi: 10.1111/j.1464-5491.2009.02732.x] [Medline: 19538239]
6. Kuhl EA, Sears SF, Conti JB. Internet-based behavioral change and psychosocial care for patients with cardiovascular disease: a review of cardiac disease-specific applications. Heart Lung 2006;35(6):374-382. [doi: 10.1016/j.hrtlng.2006.02.004] [Medline: 17137938]
7. Liu S, Dunford SD, Leung YW, Brooks D, Thomas SG, Eysenbach G, et al. Reducing blood pressure with Internet-based interventions: a meta-analysis. Can J Cardiol 2013 May;29(5):613-621. [doi: 10.1016/j.cjca.2013.02.007] [Medline: 23618507]
8. Nolan RP, Liu S, Shoemaker JK, Hachinski V, Lynn H, Mikulis DJ, et al. Therapeutic benefit of internet-based lifestyle counselling for hypertension. Can J Cardiol 2012 May;28(3):390-396. [doi: 10.1016/j.cjca.2012.02.012] [Medline: 22498181]
9. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
10. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
11. Stopponi MA, Alexander GL, McClure JB, Carroll NM, Divine GW, Calvi JH, et al. Recruitment to a randomized web-based nutritional intervention trial: characteristics of participants compared to non-participants. J Med Internet Res 2009;11(3):e38 [FREE Full text] [doi: 10.2196/jmir.1086] [Medline: 19709990]
12. 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]
13. Jeffery RW. Financial incentives and weight control. Prev Med 2012 Nov;55 Suppl:S61-S67 [FREE Full text] [doi: 10.1016/j.ypmed.2011.12.024] [Medline: 22244800]
14. Hunter RF, Tully MA, Davis M, Stevenson M, Kee F. Physical activity loyalty cards for behavior change: a quasi-experimental study. Am J Prev Med 2013 Jul;45(1):56-63. [doi: 10.1016/j.amepre.2013.02.022] [Medline: 23790989]
15. Rice T. The behavioral economics of health and health care. Annu Rev Public Health 2013;34:431-447. [doi: 10.1146/annurev-publhealth-031912-114353] [Medline: 23297657]
16. Teixeira PJ, Palmeira AL, Vansteenkiste M. The role of self-determination theory and motivational interviewing in behavioral nutrition, physical activity, and health: an introduction to the IJBNPA special series. Int J Behav Nutr Phys Act 2012;9:17 [FREE Full text] [doi: 10.1186/1479-5868-9-17] [Medline: 22385644]
17. Mitchell MS, Goodman JM, Alter DA, John LK, Oh PI, Pakosh MT, et al. Financial incentives for exercise adherence in adults: systematic review and meta-analysis. Am J Prev Med 2013 Nov;45(5):658-667. [doi: 10.1016/j.amepre.2013.06.017] [Medline: 24139781]
18. Alexander GL, Divine GW, Couper MP, McClure JB, Stopponi MA, Fortman KK, et al. Effect of incentives and mailing features on online health program enrollment. Am J Prev Med 2008 May;34(5):382-388 [FREE Full text] [doi: 10.1016/j.amepre.2008.01.028] [Medline: 18407004]
19. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994 Jan;13(1):39-46. [Medline: 8168470]
20. Miller WR, Rollnick S. Motivational Interviewing: preparing people for change. New York: Guilford Press; 2002.
21. Thomas JR, Nelson J, Silverman S. Research Methods in Physical Activity - 6th Edition. Champaign, IL: Human Kinetics; 2010.
22. Rea LM, Parker RA. Designing and conducting survey research: a comprehensive guide. San Francisco: Jossey-Bass; 1992.
23. Cohen J. The Earth is round (p < .05). American Psychologist 1994;49(12):997-1003. [doi: 10.1037/0003-066X.49.12.997]
24. Taitel MS, Haufle V, Heck D, Loeppke R, Fetterolf D. Incentives and other factors associated with employee participation in health risk assessments. J Occup Environ Med 2008 Aug;50(8):863-872. [doi: 10.1097/JOM.0b013e3181845fe2] [Medline: 18695446]

XSL•FO
RenderX

25. 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]
26. Rogers MA. Diffusion of Innovations. 5 ed. New York, NY: Free Press; 2003.
27. Charness G, Gneezy U. Incentives to exercise. Econometrica 2009;77:909-931.
28. Daley DJ, Duda JL. Self-determination, stage of readiness to change for exercise, and frequency of physical activity in young people. Eur. J. Sport Sc 2006;6:231-243. [doi: 10.1080/17461390601012637]
29. Mitchell M, Faulkner G. A "nudge" at all? The jury is still out on financial health incentives. Healthc Pap 2012;12(4):31-6; discussion 64. [Medline: 23713398]
30. Artinian NT, Fletcher GF, Mozaffarian D, Kris-Etherton P, Van Horn L, Lichtenstein AH, American Heart Association Prevention Committee of the Council on Cardiovascular Nursing. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: a scientific statement from the American Heart Association. Circulation 2010 Jul 27;122(4):406-441 [FREE Full text] [doi: 10.1161/CIR.0b013e3181e8edf1] [Medline: 20625115]

Abbreviations

CVD: cardiovascular disease
HSFC: Heart and Stroke Foundation of Canada

Long-Term Results of a Web-Based Guided Self-Help Intervention for Employees With Depressive Symptoms: Randomized Controlled Trial

Anna S Geraedts[1,2,3*], MSc; Annet M Kleiboer[1,2,3*], PhD; Jos Twisk[4,5*], PhD; Noortje M Wiezer[3,6*], PhD; Willem van Mechelen[2,3,7*], MD, PhD; Pim Cuijpers[1,2,3*], PhD

[1]Department of Clinical Psychology, Vrije University Amsterdam, Amsterdam, Netherlands

[2]EMGO Institute for Health and Care Research, VU University Amsterdam and VU University Medical Center Amsterdam, Amsterdam, Netherlands

[3]Body@Work, Research Center Physical Activity, Work and Health, TNO-VU-VUmc, Amsterdam, Netherlands

[4]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, Netherlands

[5]Department of Health Sciences, Section Methodology and Applied Biostatistics, Vrije University Amsterdam, Amsterdam, Netherlands

[6]TNO, Hoofddorp, Netherlands

[7]Department of Public and Occupational Health, VU University Medical Center, Amsterdam, Netherlands

[*]all authors contributed equally

Corresponding Author:
Anna S Geraedts, MSc
Department of Clinical Psychology
Vrije University Amsterdam
Van der Boechorststraat 1
Amsterdam, 1081 BT
Netherlands
Phone: 31 0205987451
Fax: 31 0205988758
Email: a.s.geraedts@vu.nl

Abstract

Background: Depressive disorders are highly prevalent in the working population and are associated with excessive costs. The evidence for effective worker-directed interventions for employees with depressive symptoms is limited. Treating employees with depressive symptoms via the Internet before they report sick from work could be beneficial and cost saving.

Objective: In this study, we tested the effectiveness over the period of 1 year of a Web-based guided self-help intervention, called Happy@Work, for employees with depressive symptoms who were not on sick leave.

Methods: A two-arm randomized controlled trial comparing a worker-directed, Web-based, guided self-help intervention to care as usual (CAU) was carried out. We recruited employees from 6 companies via the company's Intranet and by putting up posters. The inclusion criteria were elevated depressive symptoms as measured by a score ≥16 on the Center for Epidemiologic Studies Depression scale (CES-D) and not being on sick leave. The intervention contained 6 lessons and consisted of problem-solving treatment and cognitive therapy. Participants were asked to submit weekly assignments via the website after completion of a lesson and they received feedback from a coach via the website. Self-report questionnaires on depressive symptoms (CES-D; primary outcome), burnout (Maslach Burnout Inventory, MBI), work performance (Health and Work Performance Questionnaire, HPQ), duration of absenteeism, and anxiety (Hospital Anxiety and Depression Scale, HADS; secondary outcomes), were completed at baseline, posttreatment, and at 6-, and 12-month follow-up. Several subgroup and per-protocol analyses were performed.

Results: A total of 231 employees were randomized to either the intervention group (n=116) or to CAU (n=115). Completion of assessments varied between 54%-74%. Improvement in depressive symptoms between baseline and posttreatment was shown in all participants and these effects sustained over time. However, there were no differences between the 2 groups (adjusted regression coefficient=0.46, 95% CI –2.11 to 3.03, *P*=.72; Cohen's *d*=0.05). Differences between groups were also not significant for the secondary outcomes. No subgroups were identified to show differences between the groups, nor did we find a between-group effect in the per-protocol analyses.

XSL•FO
RenderX

Conclusions: This study showed that a worker-directed, Web-based, guided self-help intervention was not more effective than CAU in reducing depressive symptoms among employees with depressive symptoms who were not on sick leave over the period of 1 year. An intervention for this specific target group might not be necessary because the recovery in the CAU group was comparable to the intervention group and sustained over a 12-month period.

Trial Registration: Nederlands Trial Register (NTR): NTR2993; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2993 (Archived by WebCite at http://www.webcitation.org/6PL9pFC0n).

Introduction

Depressive disorders are highly prevalent in the general [1-3] and working [4,5] populations and lead to excessive costs [6,7]. Approximately 70%-85% of the costs are because of work absenteeism, work impairment, and loss of work productivity, which suggests that companies pay the largest part of the total costs of depression [8-12].

Research on the treatment of depression has been extensive and has shown that depression can be treated effectively with different forms of psychotherapies [13-18]. Traditionally, most types of psychotherapies are delivered face-to-face in mental health care settings, but there is increasing evidence for the effectiveness of treatments that are delivered via the Internet [19-23]. In general, studies on the effectiveness of Web-based interventions for the treatment of depressive symptoms show positive short-term effects [21], but there are fewer studies available that have also studied the long-term effects of Web-based interventions [21]. In a recently published meta-analysis on the effects of computer cognitive behavior therapy (CCBT) for depression, Richards and Richardson [21] reported the results of 14 studies that included a long-term follow-up, primarily up to 6 months with few studies reporting outcomes up to 12 months. They showed a small but significant effect of CCBT on depression (d=0.20) but stressed that more studies are needed to confirm the benefits of Web-based interventions at long-term follow-up [21].

The large number of studies on the treatment of depression in mental health care is in contrast with the few studies on worker-directed interventions for employees with depression or depressive symptoms. It is, however, important to develop evidence-based worker-directed interventions for employees with depression that involve work-related aspects, such as high job demands and work-life balance, because work-related aspects play an important role in the development and perpetuation of depression [24-26]. The Organisation for Economic Co-operation and Development (OECD) [4] has recently recommended to increase the evidence for worker-directed treatments of mental health problems and have highlighted the importance of intervening before employees take sick leave. Early intervention (before sick leave) is important because it may prevent worsening of mental health problems; consequently, it has the potential to reduce the costs of work absenteeism and loss of work productivity [4,26,27].

Several studies have been published on the effectiveness of face-to-face or Web-based worker-directed interventions for non-sick-listed employees [28-38]. Most of these studies were aimed at employees with stress or burnout symptoms who had not (yet) reported sick from work. All these studies showed positive effects of the interventions on symptom reduction. Care-as-usual (CAU) and waiting-list control groups were used most frequently as reference groups and the highest effects were seen in studies with a waiting-list control group. However, it is known that studies that use a waiting-list control comparator have a tendency to show stronger effect sizes of the intervention because they are less likely to positively affect the outcome compared with active control groups, such as CAU [39]. Two of these studies [30,31] examined face-to-face interventions for non-sick-listed employees with depressive symptoms. To our knowledge, no studies have been published on the effectiveness of Web-based worker-directed interventions for employees with depressive symptoms who are not on sick leave. Web-based treatments may be of special benefit to the working population because the employee will not have to take time off from work for therapist visits and participation in Web-based treatments is more anonymous compared to face-to-face treatment.

Considering the importance of developing Web-based worker-directed interventions for employees with depression and the limited knowledge on the long-term effects of such interventions, we conducted a randomized controlled trial with a long-term follow-up period of 12 months in which we examined the effects of such an intervention for employees with depressive symptoms who were not on sick leave compared to a CAU control group. The design of this study has been published elsewhere [40]. A process evaluation of this study (submitted paper) revealed that the intervention was conducted according to protocol and seemed feasible for further implementation. The posttreatment effects of the Web-based guided self-help intervention showed significant but small effect sizes in favor of the intervention group for anxiety symptoms and emotional exhaustion. The intervention group improved substantially on the primary outcome of depressive symptoms, but the CAU control group improved considerably as well and there was no significant difference between both groups [41]. It is of importance to examine whether the improvement in both groups is sustainable over time or if there will be an increase of depressive symptoms in 1 or both groups. Therefore, in this study we examined between-group differences over a 1-year follow-up period on depressive symptoms, burnout symptoms, work performance, and anxiety symptoms. In addition, we

studied the effects of the intervention on absenteeism and we performed several subgroup analyses regarding educational level, age, gender, working hours, and baseline depression score because different effects for these subgroups might be possible.

Methods

Participants

The design and short-term outcomes of this study have been described in detail elsewhere [40,41]. Therefore, we will describe the design briefly. Participants were recruited via 6 different companies in the Netherlands—2 banking companies, 2 research institutes, 1 security company, and 1 university—through banners and digital pamphlets on the company's Intranet and via posters. Employees who showed interest in the study received an information leaflet and an informed consent form via email. After participants gave informed consent, they received a link to an online screening questionnaire via email. Employees with elevated depressive symptoms as measured by a score of 16 or higher on the Center for Epidemiologic Studies Depression scale (CES-D) who were not on sick leave (at the time they completed the baseline questionnaire) were eligible to take part in the study. Furthermore, access to the Internet and an email address were required. Participants were excluded if they had been using medication for depressive symptoms for less than 1 month or if they had a legal labor dispute with the employer. Once included, participants were randomized to the Web-based intervention or the CAU control group. The recruitment and retention details are shown in Figure 1.

Figure 1. Flowchart of participants.
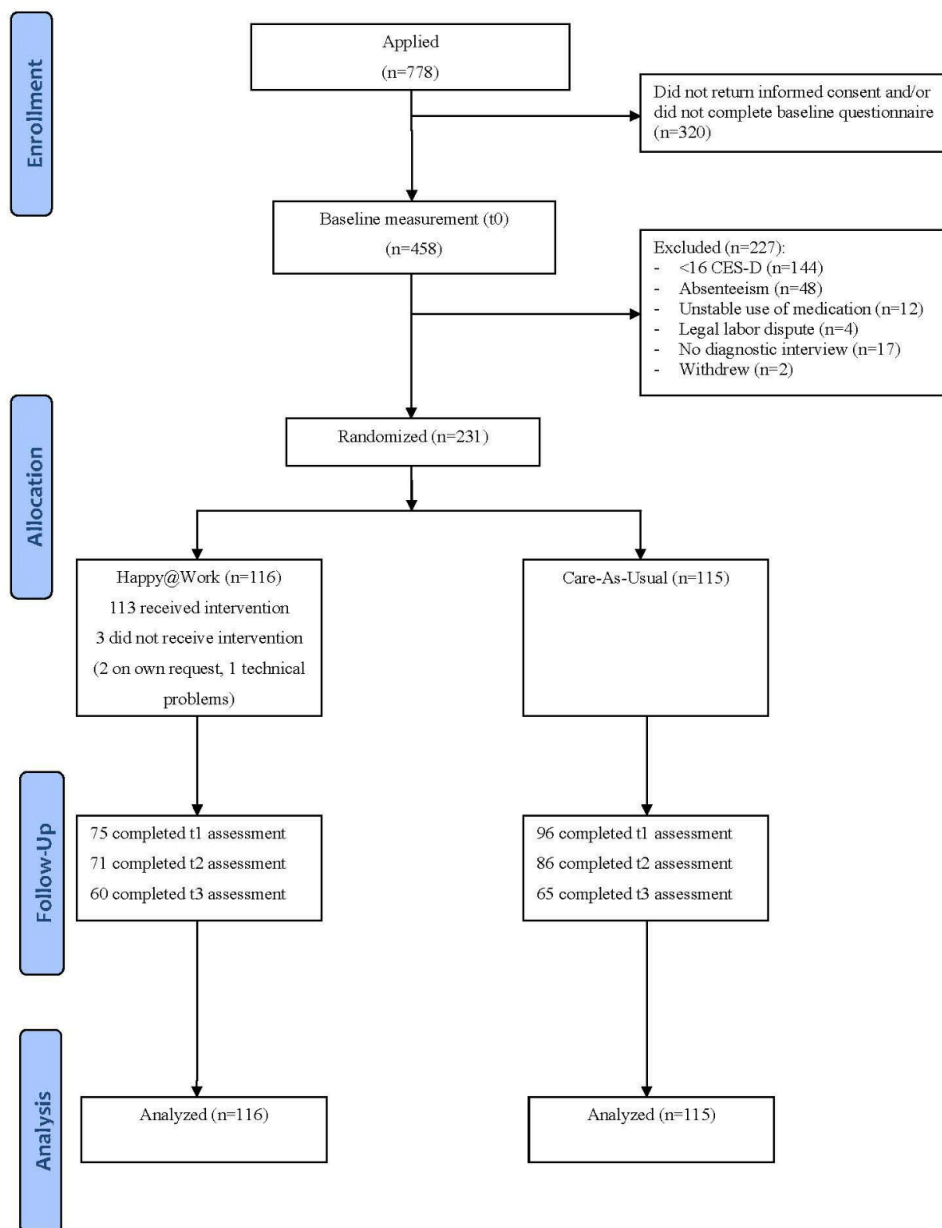
Procedure

This study was a randomized controlled trial with 2 arms: a Web-based guided self-help intervention (called Happy@Work) and a CAU group. The study was approved by the Medical Ethics Committee of the VU University Medical Center (registration number 2011/2) and registered in the Dutch Trial Register (NTR2993). The sample size was determined at 200 participants, based on a power of .80, an alpha of .05, and an expected dropout percentage of 30% to show a posttreatment effect size Cohen's d of 0.50. A total of 231 participants were randomized to the Happy@Work intervention (n=116) or the CAU group (n=115). Randomization took place at an individual level after completion of the baseline measurement (questionnaire and clinical interview). We used stratification at 2 levels: (1) use of antidepressants and (2) receiving treatment from a psychologist or psychiatrist at study entry. Block randomization was used with random blocks containing 4, 6, or 8 allocations. An independent researcher made the allocation schedule with a computerized random number generator and the investigators had no knowledge of the schedule. The participants were informed about randomization outcome via email. Participants completed online questionnaires at baseline and posttreatment at 8 weeks (t1), 6 months (t2), and 12 months (t3).

Interventions

Happy@Work

The intervention Happy@Work [42] is a brief Web-based intervention delivered with minimal guidance. It consists of 2 evidence-based treatments; problem-solving treatment (PST)

[43] and cognitive therapy (CT) [44], and a guideline for employees to help them to prevent work-related stress [45,46]. Happy@Work consists of 6 weekly lessons with an option of 1 week extra time in case of delay. Each lesson has a different theme, but always follows the same structure: information about the theme, examples, and assignments. Themes of the lessons are introduction of problem solving (lesson 1), problem-solving methods (lesson 2), changing cognitions (lesson 3), dealing with work-related problems (lesson 4), social support (lesson 5), and relapse prevention (lesson 6). Participants receive feedback on assignments from a coach. Coaches were trained Master's-level students in clinical psychology. All coaches used a protocol-treatment manual. To ensure treatment fidelity, all feedback was reviewed by a supervisor (AG) before it was placed on the website. Happy@Work is a tunneled intervention, which means that participants can start with a new lesson after they have received feedback on their assignments from a coach. Participants were viewed as treatment completers if they had followed at least the basic information and assignments of PST and CT (completion of lessons 1-3).

At the start of the intervention, an account was generated by the researchers on the website and a coach was assigned to the participant on the website. Once the account was generated, an automatic email was sent to the participant with a link to activate the account. Participants used their email address and a self-created password to log in once the account was activated. Reminders were sent to participants via email when deadlines were not met. There were no changes to the content, bugs, or periods with downtime during the trial. Screenshots of the intervention can be found in Figure 2 and in Multimedia Appendix 1.

Figure 2. Screenshot of the Happy@Work intervention.
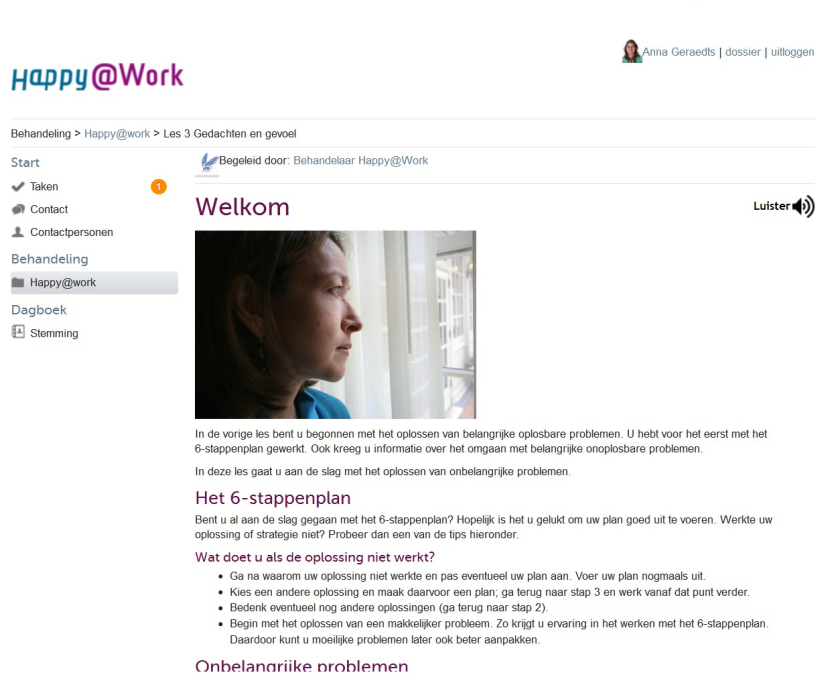
Care as Usual

Participants randomized to the CAU group received an email with the randomization outcome only and were advised to consult their (occupational) physician or a psychologist if they wanted treatment for their depressive symptoms. Participants in both conditions were free to seek any additional (mental) health care.

Measures

Depressive Symptoms

The primary outcome was depressive symptoms as measured by the CES-D [47]. This questionnaire is widely used for identifying people with depressive symptoms. Its validity has been tested in different populations [48-50]. The CES-D consists of 20 items and the total score varies between 0 and 60. The baseline Cronbach alpha in this study was .82. A score of 16 or higher represents a clinically significant level of depressive symptoms [47]. The cut-off score of 16 was used in this study as an inclusion criterion. This cut-off score is used frequently in studies and has shown to have good sensitivity (0.95), specificity (0.85), and positive predictive value of major depression (0.11) in a sample of employees [50].

Burnout Symptoms

Burnout symptoms were measured with the Dutch version of the Maslach Burnout Inventory-General Scale (MBI) [51,52]. This self-report questionnaire contains 15 items and 3 dimensions: emotional exhaustion (5 items), cynicism (4 items), and reduced professional efficacy (6 items). Every item was scored on a 7-point Likert scale (0-6). Following the manual of the questionnaire [52], a total score for every dimension was calculated by adding the item scores and by dividing the total score by the number of items, with higher scores indicating more severe symptoms. We rescored the professional efficacy dimension with higher scores indicating less feeling of professional efficacy. The baseline Cronbach alphas for the different dimensions in this study were .83 for exhaustion, .83 for cynicism, and .79 for reduced professional efficacy.

Work Absenteeism

Work absenteeism was measured with the second part of the Trimbos and iMTA Questionnaire on Costs Associated with Psychiatric Illness (TiC-P) the Short Form Health and Labor Questionnaire (SF-HLQ) [53]. The participant was asked to report the total number of days absent from work because of illness in the time period between the assessments at 8 weeks (t0-t1), 4 months (t1-t2), and 6 months (t2-t3). The recall period at baseline assessment was 3 months. Research has shown that participants can report valid and accurate rates of work absenteeism up to 6 months [54].

Work Performance

We used the general work performance scale of the World Health Organization (WHO) Health and Work Performance Questionnaire (HPQ) [55], which contains 4 items. Item 4 gives the best and easiest indication of the participant's perception of their own work performance [56] by asking participants to rate their overall work performance during the past 4 weeks compared to employees in comparable functions. We only report

on that item in this study. Work performance was scored on a 7-point Likert scale with a higher score indicating poorer work performance compared to other employees [56].

Anxiety Symptoms

The anxiety subscale of the Hospital Anxiety and Depression Scale (HADS) was used to measure anxiety symptoms [57]. The anxiety subscale of the HADS consists of 7 items. Scores range from 0 to 21, with higher scores indicating more anxiety. The HADS has shown good homogeneity and reliability in different normal and clinical Dutch samples [58]. The baseline Cronbach alpha in this study was .76.

Clinical Interview

The WHO Composite International Diagnostic Interview version 2.1 (CIDI) [59] is a structured interview to assess psychiatric diagnosis defined in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, Text Revision (*DSM-IV-TR*) [60]. For this study, 2 sections of the CIDI were assessed: the mood disorders section and the "other" anxiety disorders (social phobia, panic disorder, agoraphobia, and generalized anxiety disorder) section. The CIDI was conducted by trained interviewers via telephone at baseline (T0) and 6-month follow-up and was used for diagnostic purposes.

Health Care Utilization

A revised version of the Trimbos and iMTA Questionnaire on Costs Associated with Psychiatric Illness (TiC-P) [53] was used to collect data on health care utilization. The TiC-P is a self-report questionnaire and consists of 2 different parts that can be administered separately. Part I was used, which contains 12 items concerning health care utilization by participants. There were 2 questions added to the questionnaire about the frequency of utilization of different health care services of the company: occupational physician and occupational social work. The questionnaire was used at T0 to assess health care utilization up to 3 months before the start of the study and at posttreatment (t1) assessment to assess health care utilization between baseline and posttreatment assessment.

Other Measures

We included several demographic questions and questions about working hours and working days in the baseline questionnaire.

Statistical Analyses

Effectiveness

Linear mixed modeling (LMM) was used to examine treatment differences. Two LMM analyses were performed: (1) unadjusted analyses, only controlling for the baseline score of the outcome measure and (2) adjusted analyses, controlling for other baseline variables, such as age, gender, marital status, educational level, nationality, and working hours, as well as the baseline outcome score. In LMM analyses, the regression coefficient represents the overall mean difference between the groups over time, so over all assessments after baseline. Reporting the overall mean difference between the groups over time was chosen because we were interested in the difference between the groups over the entire period of 1 year. If the regression coefficient is

positive, the mean difference is in favor of the intervention group; if the regression coefficient is negative, the mean difference is in favor of the CAU group.

An overall between-group effect size for every outcome variable was calculated according to Cohen's *d* [61]. The Cohen's *d* was calculated by dividing the overall mean difference between the groups (expressed as regression coefficient) by the overall SD of the observed data. Effect sizes ≥0.8 are assumed to be large, effect sizes between 0.5-0.8 are moderate, and effect sizes between 0.2-0.5 are assumed to be small [61]. Furthermore, in additional analyses we calculated the Cohen's *d* for depressive symptoms on every assessment based on the observed data. The Cohen's *d* was calculated by subtracting the mean score of the CAU group from the mean score of the intervention group and dividing that result by the pooled standard deviation.

All analyses were performed according to the intention-to-treat (ITT) principle. Missing data were handled by multiple imputation via data augmentation. Data augmentation is an iterative Markov chain Monte Carlo method to generate the imputed values assuming a multivariate normal distribution. Five imputations were used in all analyses and reported in the effectiveness analyses. Results of the mean and standard deviations reported are of the observed data.

Subgroup and Per-Protocol Analyses

We performed several a priori subgroup analyses on the primary outcome depressive symptoms. These subgroup analyses were educational level, age (age <35 versus age ≥35), gender, working full time (≥36 hours per week) versus part time (<36 hours per week), and high baseline score as defined by a score of ≥27 on the CES-D (used more often as an indicator of more severe depressive symptoms) [62-64]. In the subgroup analyses, the specific subgroup was selected from all study participants and the difference between the groups over time was then compared.

Furthermore, we performed a per-protocol analysis based on treatment completers (completed ≥3 lessons of the intervention).

Sensitivity Analyses

We also performed all analyses on the data for 100 imputations. Because there is a current debate whether it is necessary to perform multiple imputations in combination with mixed-model analyses in longitudinal studies [65], we also performed the LMM analyses without multiple imputations. All multiple imputations and LMM analyses were performed in STATA version 11.2 (StataCorp LP, College Station, TX, USA) with the procedures mi and xtmixed.

Results

Participants and Response Rates

Figure 1 shows the flow of participants through the trial. A total of 231 participants were included in the trial, 29.7% (231/778) of the employees who showed initial interest in the study. Of these, 116 participants were randomized to the intervention group and 115 to the CAU group. Most participants (n=166) were employed by 1 of the 2 banking companies, 39 by the 2 research institutes, 11 by the security company, and 15 by the university. Of the 231 participants, 10 (4.3%) used medication without psychological treatment, 24 (10.4%) received psychological treatment but no medication, and 4 participants (1.7%) used both medication and received psychological treatment at baseline. Thus, most participants in both groups (83.6%, 193/231) were not receiving treatment for their depressive symptoms at baseline.

As shown in Table 1, most participants were female (62.3%, 144/231), born in the Netherlands (95.2%, 220/231), involved in an intimate relationship (76.2%, 176/231), highly educated (63.6%, 147/231), and worked for 34 hours per week on average.

Table 1. Participants' demographic characteristics at baseline.

Characteristic	All (N=231)	Intervention (n=116)	CAU (n=115)	P
Age (years), mean (SD)	43.4 (9.2)	43 (8.9)	43.8 (9.6)	.51
Gender, n (%)				.20
Female	144 (62.3)	77 (66.4)	67 (58.3)	
Male	87 (37.7)	39 (33.6)	48 (41.7)	
Country of birth, n (%)				.03
Netherlands	220 (95.2)	107 (92.2)	113 (98.3)	
Other	11 (4.8)	9 (7.8)	2 (1.7)	
Marital status, n (%)				.46
Relationship	176 (76.2)	86 (74.1)	90 (78.3)	
No relationship	55 (23.8)	30 (25.9)	25 (21.7)	
Education, [a] n (%)				.25
Low	16 (6.9)	11 (9.5)	5 (4.3)	
Middle	68 (29.4)	31 (26.7)	37 (32.2)	
High	147 (63.6)	74 (63.8)	73 (63.5)	
Working hours,[b] mean (SD)	33.9 (5.0)	33.7 (4.8)	34.0 (5.3)	.65
Working days, mean (SD)	4.3 (0.7)	4.3 (0.6)	4.2 (0.7)	.32

[a]Low: primary education or lower general secondary education; middle: intermediate vocational education or high school; high: higher vocational education or university.

[b]Mean working hours per week according to contract of the employee.

Diagnosis

All participants completed the baseline clinical interview. At 6-month follow-up, a total of 170 participants (73.6%, 170/231) participated in the clinical interview. A total of 57 participants (24.7%) were diagnosed with a current major depressive disorder, dysthymic disorder, or both at baseline: 23 participants from the intervention group and 34 in the CAU group. At 6-month follow-up, 19 participants were diagnosed with a current major depressive disorder, dysthymic disorder, or both: 6 participants from the intervention group and 13 in the CAU group. From the 57 participants who were diagnosed with a current major depressive disorder, dysthymic disorder, or both at baseline, 9 participants suffered from a current major depressive disorder, dysthymic disorder, or both at 6-month follow-up as well: 2 participants from the intervention group and 7 participants from the CAU group. There were 10 participants who were diagnosed with a current major depressive disorder, dysthymic disorder, or both at 6-month follow-up but not at baseline. Of those 10 participants, 4 participants were from the intervention group and 6 participants were from the CAU group.

Health Care Utilization

At posttreatment, we analyzed the health care utilization of both groups to get a more detailed view on health care utilization by the CAU group. Only a small number of the total participants made use of health care and this was not significantly different between the groups. A detailed description of health care use can be found elsewhere [41].

Attrition and Adherence

Study Attrition

The attrition rates for the study sample were 26% at posttreatment assessment, 32% at the 6-month follow-up assessment, and 46% at the 12-month follow-up assessment. Participants in the CAU group completed the posttreatment assessment (χ^2_1=11.5, P=.001) and the 6-month follow-up assessment (χ^2_1=4.9, P=.03) more often. There were no differences between the groups for completion of the 12-month follow-up assessment. Attrition rates for the posttreatment assessment were lower in participants who completed the intervention (χ^2_1=32.1, P<.001).

Intervention Adherence

Of the 116 participants randomized to the intervention group, 9.5% (11/116) did not start or complete the first lesson of Happy@Work. A total of 67 participants (57.8%) were seen as treatment completers because they completed 3 or more lessons of the intervention. A total of 29 of 116 participants dropped out of the intervention at their own request or because of prolonged inactivity on the website. The other participants were not able to complete more lessons within the time limit of 7 weeks. Most participants who dropped out did not report a reason for dropout (15/116, 12.9%). When reasons were reported (14/116), they pertained mostly to lack of time (8/14, 57.1%).

XSL•FO
RenderX

Table 2. Observed scores of the intervention and care-as-usual (CAU) groups on different outcome measures.

Outcome	Assessment time, mean (SD)			
	Baseline (t0) (n=231)	Posttreatment (t1) (n=171)	Follow-up 6 months (t2) (n=157)	Follow-up 12 months (t3) (n=125)
CES-D				
Intervention	25.7 (7.5)	15.8 (10.6)	15.7 (11.3)	13.8 (9.7)
CAU	26.1 (7.0)	18.3 (9.1)	14.5 (8.9)	16.2 (10.7)
MBI-exhaustion				
Intervention	3.3 (1.2)	2.7 (1.2)	2.6 (1.4)	2.3 (1.4)
CAU	3.3 (1.1)	3.0 (1.2)	2.5 (1.2)	2.5 (1.3)
MBI-cynicism				
Intervention	2.8 (1.3)	2.4 (1.3)	2.5 (1.5)	2.4 (1.4)
CAU	3.1 (1.3)	2.8 (1.3)	2.4 (1.3)	2.4 (1.4)
MBI-reduced professional efficacy				
Intervention	2.6 (1.0)	2.4 (1.0)	2.3 (1.1)	2.2 (1.2)
CAU	2.7 (0.9)	2.5 (0.9)	2.3 (0.9)	2.3 (1.1)
Absenteeism (days) [a]				
Intervention	1.8 (2.7)	0.4 (1.0)	3.6 (9.4)	7.3 (25.6)
CAU	2.0 (3.3)	1.6 (4.9)	5.0 (13.7)	6.9 (23.3)
Work performance				
Intervention	4.1 (1.6)	3.6 (1.5)	3.6 (1.5)	3.6 (1.5)
CAU	4.3 (1.8)	3.6 (1.5)	3.6 (1.5)	3.7 (1.6)
HADS				
Intervention	10.6 (3.8)	7.6 (3.8)	6.8 (4.1)	6.6 (4.1)
CAU	10.2 (3.2)	8.3 (3.6)	7.2 (4.0)	6.8 (4.0)

[a]Recall periods differed per assessment: 3 months (t0), 8 weeks (t1), 4 months (t2), 6 months (t3).

Effectiveness

All participants improved between baseline and posttreatment on the primary outcome depressive symptoms and this improvement sustained over time (see Table 2). However, the overall estimated mean difference between the groups over time was not significant (see Table 3). This indicates that the estimated mean difference between the groups over the period of 1 year was not significant. The overall between-group effect size was small (d=0.05). The Cohen's d per assessment were all small to moderate effect sizes and nonsignificant (t1: d=0.26, 95% CI –0.04 to 0.56; t2: d=–0.12, 95% CI –0.43 to 0.20; t3: d=0.24 95% CI –0.12 to 0.59).

For the secondary outcomes, the same pattern of results was seen as with the depressive symptoms. There were improvements between baseline and posttreatment assessment on the secondary outcomes and these improvements sustained over time (see Table 2), but there were no significant differences between the groups over time. The overall between-group effect sizes for the secondary outcomes were all small (see Table 3). The absenteeism outcome was expressed in duration of absenteeism during the time period between 2 assessments. Therefore, it is not possible to study whether there was an increase or decrease of absenteeism duration over time, but only the differences between the groups on absenteeism duration can be examined. The overall estimated mean difference between the groups over time was not significant (see Table 3).

Table 3. Overall effectiveness on different outcome measures.

Outcome	Unadjusted coefficient[a]	95% CI	P	Effect size[b]	Adjusted coefficient[c]	95% CI	P	Effect size[b]
CES-D	0.14	–2.00, 2.27	.90	0.01	0.46	–2.11, 3.03	.72	0.05
MBI-exhaustion	0.10	–0.14, 0.33	.42	0.08	0.10	–0.13, 0.33	.40	0.08
MBI-cynicism	–0.08	–0.33, 0.17	.54	–0.06	–0.07	–0.32, 0.18	.57	–0.05
MBI-reduced professional efficacy	0.00	–0.24, 0.24	.98	0.00	0.04	–0.20, 0.27	.76	0.04
Absenteeism	–0.01	–4.69, 4.67	.99	0.00	–0.89	–6.09, 4.31	.72	0.04
Work performance	0.05	–0.24, 0.35	.72	0.03	0.01	–0.30, 0.32	.94	0.01
HADS	0.48	–0.29, 1.25	.22	0.12	0.60	–0.19, 1.38	.13	0.15

[a]Unadjusted regression coefficient: analyses adjusted for baseline outcome score.

[b]The effect size is presented as an overall effect size represented as Cohen's *d*: the number of standard deviations in the intervention group has improved more than the CAU group.

[c]Adjusted regression coefficient: analyses adjusted for baseline variables and baseline outcome score.

Subgroup Analyses

Data from the a priori subgroup analyses are reported in Table 4. There were no significant differences in depressive symptoms between the groups over time in any of the subgroups. Because the coefficients from the different subgroups were not substantially different from each other, there were no additional interaction effects tested to study whether there was a difference between the different subgroups over time.

Table 4. Overall effectiveness on depressive symptoms in different subgroups.

Subgroup	Unadjusted coefficient[a]	95% CI	P
Gender			
Female	0.60	–2.13, 3.33	.66
Male	–0.35	–4.06, 3.37	.85
Educational level			
Low	–0.24	–11.95, 11.46	.97
Middle	1.12	–3.30, 5.53	.61
High	–0.34	–2.89, 2.21	.80
Baseline CES-D score			
Score <27	0.76	–2.05, 3.60	.59
Score ≥27	–0.37	–4.62, 3.89	.86
Age			
Age <35	–0.22	–5.10, 4.66	.93
Age ≥35	0.28	–2.05, 2.60	.82
Working hours			
Work part time	–0.95	–4.05, 2.16	.55
Work full time	0.93	–1.95, 3.82	.52

[a]Unadjusted regression coefficient: analyses adjusted for baseline depression score.

Per-Protocol Analyses

The per-protocol analyses, in which the group of treatment completers was compared to the CAU group, did not reveal any significant results on the primary outcome depressive symptoms (unadjusted regression coefficient=–0.48, 95% CI –4.28 to 3.33, *P*=.80) and all secondary outcomes. The overall estimated mean difference for the MBI exhaustion dimension was 0.10 (95% CI –0.24 to 0.43, *P*=.57), for the MBI cynicism dimension it was 0.14 (95% CI –0.21 to 0.49, *P*=.42), for the MBI reduced professional efficacy dimension it was –0.03 (95% CI –0.48 to 0.41, *P*=.88), for work performance it was –0.14 (95% CI –0.79 to 0.51, *P*=.65), for absenteeism it was –1.66 (95% CI –7.10 to 3.78, *P*=.54), and for anxiety symptoms it was 0.08 (95% CI –1.06 to 1.23, *P*=.89).

Sensitivity Analyses

The analyses from the datasets without imputations and with 100 imputations did not reveal any relevant differences compared to the outcomes from the dataset with 5 imputations (data not shown).

Discussion

Principal Results

This study examined the long-term effects of a worker-directed, Web-based, guided self-help intervention on depressive symptoms, several work-related outcome measures, and anxiety symptoms compared to CAU in employees with depressive symptoms who were not on sick leave. This study did not affirm evidence for the long-term effectiveness of the Web-based intervention compared to CAU for any of the outcome measures. Overall, participants improved substantially on the primary outcome depressive symptoms between baseline and posttreatment assessment and these improvements sustained over the period of 1 year. This was also true for the work-related outcomes of burnout symptoms and work performance as participants improved between baseline and posttreatment with sustainable effects up to 12 months. Overall, participants further improved after posttreatment assessment on anxiety symptoms. However, no difference between the 2 conditions in the course of symptoms was found on any of the outcome measures. Furthermore, there were no significant mean differences between the groups on duration of absenteeism during the follow-up period. We were not able to identify any subgroups that benefited from the treatment compared to CAU. Participants with a relatively high or low score on depressive symptoms, male or female, age <35 or >35 years, working part time or full time, having a low, middle, or high educational level, or who had completed treatment or not did not improve more than the CAU group with respect to their depressive symptoms.

Comparison With Prior Work

The results of this study regarding depressive symptoms are not in line with the positive findings of the meta-analysis on the long-term effects of CCBT for depression by Richards and Richardson [21]. The Cohen's ds that were assessed at each time point, based on the observed data, showed effect sizes that were close to the overall effect size of Richards and Richardson, but they were not significant and the effect size over time was small (d=0.05). There are 2 important differences between the studies analyzed in the meta-analysis and our study which make the results of the meta-analysis more difficult to compare to this study: (1) in general, the studies in the meta-analysis examined a target group with more severe depressive symptoms and/or depressive disorders at baseline compared to this study, and (2) none of the studies in the meta-analysis were tested in a workplace context. Two studies have been published on the effects of Web-based interventions in a workplace context that included long-term follow-up results and that focused on a comparable target group of non-sick-listed employees with mild to moderate depressive symptoms at baseline. Both studies tested unguided Web-based interventions. One of these interventions was a worker-directed intervention [28] and the other had no specific focus on work-related problems [66]. Both

studies did not report significant effects in favor of the intervention at follow-up (either 3 or 6 months) and showed the same pattern of improvement as was found in this study: substantial improvements between baseline and posttreatment which sustained at follow-up in both groups. This pattern of improvement was also seen in this study, but not in other studies with long-term follow-up assessments [21].

The large reduction in depressive symptoms in the CAU group between baseline and posttreatment was unforeseen and sustained at the follow-up assessments [34,67,68]. We discussed several potential reasons for the large reduction of depressive symptoms in the CAU group when we reported the posttreatment effectiveness of this study. These were spontaneous recovery, a phenomenon which is seen more often in patients with depression [69], recruitment of highly motivated employees who were willing to change which could have let to improvement by itself, positive influences of work (ie, being able to function and stay at work while experiencing depressive symptoms might have had a positive influence on recovery of depressive symptoms), a company's participation in this study gives a positive signal of an open environment to employees (ie, a change in organizational culture) which could have led to participants in the CAU group discussing their mental health problems with their supervisor which can result in reduction of depressive symptoms, and the email with randomization outcome for the CAU group contained advice to seek treatment for depressive symptoms. This email could have instigated a behavioral change according to the stages-of-change model from Prochaska and colleagues [70]. Only a small percentage of the participants in the CAU group reported having received professional help. However, it could be possible that other participants received help in a different way; for example, via their significant other or via other self-help treatments. In relation to the latter reason for the reduction of symptoms in the CAU group, it could also be possible that for this specific target group filling in a questionnaire about depression during a period of sad mood could have been enough of an intervention by itself. Considering the comparable pattern of findings of this study and the study of Grime [28] and Phillips and colleagues [66] in non-sick-listed employees, it may be possible that spontaneous recovery of depressive symptoms is more common in this specific target group, but all these reasons could have contributed to the large improvements in the control group.

When we examined the posttreatment effects of the Web-based guided self-help intervention on burnout symptoms [41], we found small but significant differences in favor of the intervention group for emotional exhaustion but not on the other 2 dimensions, cynicism and reduced professional efficacy. We explained this finding by postulating that a change in emotional exhaustion might show a first indication of treatment effect on burnout and that the other dimensions, cynicism and reduced professional efficacy, would follow because these are related to cognitions and attitudes that generally take a longer time to show improvement. Apparently, this was not the case because no further improvements on the cynicism and reduced professional efficacy dimensions occurred at follow-up, but the small improvements between baseline and posttreatment assessment sustained during follow-up.

To our knowledge, this is the first study on Web-based interventions that has used absenteeism as an outcome measure. We did not find between-group differences in absenteeism, but we were not able to investigate if there was an increase or decrease in absenteeism over time because of the use of different time periods between assessments. Future research on Web-based interventions, especially when tested in the workplace context, should include absenteeism duration and frequency as an outcome measure.

Limitations

This study has several limitations. The first has to do with the attrition rate and handling of missing data. We were confronted with a high attrition rate which is seen more often in Web-based interventions [71,72]. The attrition rates in this study were equal or lower compared with several similar studies on guided Web-based interventions for depression with long-term follow-up assessments [67,73,74]. The bias that may have been introduced was accounted for by applying multiple imputation techniques. Because of the current debate on the necessity of multiple imputations in combination with mixed-model analysis in longitudinal studies [65], we also performed mixed-model analysis without multiple imputations. The results were comparable, indicating that data were robust and multiple imputations may not have been needed. Second, the participants in this study were primarily Dutch white-collar workers with a high educational level. Therefore, it is uncertain whether the results can be generalized to the general working population or employees with a lower education level. Although our subgroup analysis on educational level did not show significant differences, the subgroup analyses had a lack of power and only 36.4% of the study population had a low or middle educational level. Third, the power-analysis was based on a posttreatment effect and, therefore, the analyses on follow-up assessments have a lack of power. Finally, as stated previously, adherence to the intervention was low and only 57.8% completed at least 3 lessons of the intervention. Therefore, the analyses of comparisons between the intervention group and the CAU group compared the effects of a low adherence intervention, with many participants who only followed a small part of the intervention. The per-protocol analyses did not show significant differences

either, but had a lack of power because the analyses were only based on 42.2% of the intervention group.

Implications and Future Research

The results of this study implicate that the intervention Happy@Work is not more effective in reducing depressive symptoms than CAU over the period of 1 year. Overall, participants improved substantially between baseline and posttreatment assessment on depressive symptoms and these improvements sustained over time. Participants also improved on the secondary outcomes, which sustained over time. The large improvements on depressive symptoms in the CAU group were also found in 2 studies with comparable target groups of non-sick-listed employees [28,66]. Therefore, it could be possible that spontaneous recovery of depressive symptoms is more likely in this specific target group. Observational research following non-sick-listed depressed employees over time could provide more insight.

The process evaluation that was performed alongside this trial concluded that the intervention was feasible for further implementation. However, based on the results of this trial we do not recommend to directly implement Happy@Work into routine practice because it was not more effective than CAU over time. It could, however, be possible that the intervention, even though it is not effective from a clinical perspective, could be effective from an economical perspective (eg, cost-effective). This needs further investigation. Further, more research is needed to examine the possibilities of using e-mental health in the workplace setting. This research should focus on the needs of employees with mental health problems and on the ideal moment when intervention is really necessary.

Conclusions

This study showed that the Web-based, worker-directed, guided self-help intervention Happy@Work is not more effective in reducing depressive symptoms than a CAU group over the period of 1 year. Based on the results of this study, we can conclude that an intervention for employees with mild to moderate depressive symptoms who are not on sick leave might not be necessary because the natural recovery in the CAU group was comparable to the intervention group and sustainable over a 12-month period.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the intervention.

[PDF File (Adobe PDF File), 241KB - jmir_v16i7e168_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [75].

[PDF File (Adobe PDF File), 989KB - jmir_v16i7e168_app2.pdf]

References

1. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004(420):21-27. [doi: 10.1111/j.1600-0047.2004.00327.x] [Medline: 15128384]

2. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994 Jan;51(1):8-19. [Medline: 8279933]

3. Waraich P, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature. Can J Psychiatry 2004 Feb;49(2):124-138. [Medline: 15065747]

4. Organisation for Economic Co-operation and Development. Sick on the job? Myths and realities about mental health and work. OECD Publishing 2011 Dec 01:2012. [doi: 10.1787/9789264124523-en]

5. Wang J, Adair CE, Patten SB. Mental health and related disability among workers: a population-based study. Am J Ind Med 2006 Jul;49(7):514-522. [doi: 10.1002/ajim.20319] [Medline: 16634079]

6. Berto P, D'Ilario D, Ruffo P, Di Virgilio R, Rizzo F. Depression: cost-of-illness studies in the international literature, a review. J Ment Health Policy Econ 2000 Mar 1;3(1):3-10. [Medline: 11967432]

7. Greenberg PE, Birnbaum HG. The economic burden of depression in the US: societal and patient perspectives. Expert Opin Pharmacother 2005 Mar;6(3):369-376. [doi: 10.1517/14656566.6.3.369] [Medline: 15794728]

8. de Graaf R, Tuithof M, van Dorsselaer S, ten Have M. Verzuim door psychische en somatische aandoeningen bij werkenden. Resultaten van de 'Netherlands Mental Health Survey and Incidence Study-2' (NEMESIS-2) [Absenteeism due to psychological or somatic disorders in workers. Results of the 'Netherlands Mental Health Survey and Incidence Study-2' (NEMESIS-2)]. Utrecht: Trimbos-Instituut; 2011.

9. Henderson M, Glozier N, Holland Elliott K. Long term sickness absence. BMJ 2005 Apr 9;330(7495):802-803 [FREE Full text] [doi: 10.1136/bmj.330.7495.802] [Medline: 15817531]

10. Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? J Occup Environ Med 2008 Apr;50(4):401-410. [doi: 10.1097/JOM.0b013e31816bae50] [Medline: 18404013]

11. Smit F, Cuijpers P, Oostenbrink J, Batelaan N, de Graaf R, Beekman A. Costs of nine common mental disorders: implications for curative and preventive psychiatry. J Ment Health Policy Econ 2006 Dec;9(4):193-200. [Medline: 17200596]

12. Thomas CM, Morris S. Cost of depression among adults in England in 2000. Br J Psychiatry 2003 Dec;183:514-519 [FREE Full text] [Medline: 14645022]

13. Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, et al. A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. Health Technol Assess 2001;5(35):1-173 [FREE Full text] [Medline: 12387733]

14. Cuijpers P, Dekker J, Hollon SD, Andersson G. Adding psychotherapy to pharmacotherapy in the treatment of depressive disorders in adults: a meta-analysis. J Clin Psychiatry 2009 Sep;70(9):1219-1229. [doi: 10.4088/JCP.09r05021] [Medline: 19818243]

15. Cuijpers P, Geraedts AS, van Oppen P, Andersson G, Markowitz JC, van Straten A. Interpersonal psychotherapy for depression: a meta-analysis. Am J Psychiatry 2011 Jun;168(6):581-592 [FREE Full text] [doi: 10.1176/appi.ajp.2010.10101411] [Medline: 21362740]

16. Cuijpers P, van Straten A, Warmerdam L. Problem solving therapies for depression: a meta-analysis. Eur Psychiatry 2007 Jan;22(1):9-15. [doi: 10.1016/j.eurpsy.2006.11.001] [Medline: 17194572]

17. Dobson KS. A meta-analysis of the efficacy of cognitive therapy for depression. J Consult Clin Psychol 1989 Jun;57(3):414-419. [Medline: 2738214]

18. Pinquart M, Duberstein PR, Lyness JM. Treatments for later-life depressive conditions: a meta-analytic comparison of pharmacotherapy and psychotherapy. Am J Psychiatry 2006 Sep;163(9):1493-1501. [doi: 10.1176/appi.ajp.163.9.1493] [Medline: 16946172]

19. 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]

20. Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS Spectr 2009 Feb;14(2 Suppl 3):34-40. [Medline: 19238128]

21. 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]

22. 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]

23. Foroushani PS, Schneider J, Assareh N. Meta-review of the effectiveness of computerised CBT in treating depression. BMC Psychiatry 2011;11:131 [FREE Full text] [doi: 10.1186/1471-244X-11-131] [Medline: 21838902]

24. Bonde JP. Psychosocial factors at work and risk of depression: a systematic review of the epidemiological evidence. Occup Environ Med 2008 Jul;65(7):438-445. [doi: 10.1136/oem.2007.038430] [Medline: 18417557]

25. Nieuwenhuijsen K, Bültmann U, Neumeyer-Gromen A, Verhoeven AC, Verbeek JH, van der Feltz-Cornelis CM. Interventions to improve occupational health in depressed people. Cochrane Database Syst Rev 2008 Apr(2):CD006237. [doi: 10.1002/14651858.CD006237.pub2] [Medline: 18425942]

26. Szeto AC, Dobson KS. Mental disorders and their association with perceived work stress: an investigation of the 2010 Canadian Community Health Survey. J Occup Health Psychol 2013 Apr;18(2):191-197. [doi: 10.1037/a0031806] [Medline: 23458060]

27. Järvisalo J, Andersson B, Boedeker W, Houtman I. Mental Disorders as a Major Challenge in Prevention of Work Disability: Experiences in Finland, Germany, the Netherlands and Sweden. Helsinki: Social Insurance Institution, Research department; 2005.

28. Grime PR. Computerized cognitive behavioural therapy at work: a randomized controlled trial in employees with recent stress-related absenteeism. Occup Med (Lond) 2004 Aug;54(5):353-359 [FREE Full text] [doi: 10.1093/occmed/kqh077] [Medline: 15289593]

29. Hasson D, 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;5:78 [FREE Full text] [doi: 10.1186/1471-2458-5-78] [Medline: 16042796]

30. Lerner D, Adler D, Hermann RC, Chang H, Ludman EJ, Greenhill A, et al. Impact of a work-focused intervention on the productivity and symptoms of employees with depression. J Occup Environ Med 2012 Feb;54(2):128-135 [FREE Full text] [doi: 10.1097/JOM.0b013e31824409d8] [Medline: 22252528]

31. Lexis MA, Jansen NW, Huibers MJ, van Amelsvoort LG, Berkouwer A, Tjin A Ton G, et al. Prevention of long-term sickness absence and major depression in high-risk employees: a randomised controlled trial. Occup Environ Med 2011 Jun;68(6):400-407. [doi: 10.1136/oem.2010.057877] [Medline: 20924024]

32. Van Rhenen W, Blonk RW, van der Klink JJ, van Dijk FJ, Schaufeli WB. The effect of a cognitive and a physical stress-reducing programme on psychological complaints. Int Arch Occup Environ Health 2005 Mar;78(2):139-148. [doi: 10.1007/s00420-004-0566-6] [Medline: 15761748]

33. van Rhenen W, Blonk RW, Schaufeli WB, van Dijk FJ. Can sickness absence be reduced by stress reduction programs: on the effectiveness of two approaches. Int Arch Occup Environ Health 2007 May;80(6):505-515. [doi: 10.1007/s00420-006-0157-9] [Medline: 17093962]

34. 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]

35. Willert MV, Thulstrup AM, Hertz J. Changes in stress and coping from a randomized controlled trial of a three-month stress management intervention. Scand J Work Environ Health 2009 Mar;35(2):145-152 [FREE Full text] [Medline: 19308298]

36. Willert MV, Thulstrup AM, Bonde JP. Effects of a stress management intervention on absenteeism and return to work--results from a randomized wait-list controlled trial. Scand J Work Environ Health 2011 May;37(3):186-195 [FREE Full text] [Medline: 21057736]

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. 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]

39. Mohr DC, Spring B, Freedland KE, Beckner V, Arean P, Hollon SD, et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. Psychother Psychosom 2009;78(5):275-284. [doi: 10.1159/000228248] [Medline: 19602916]

40. Geraedts AS, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P. Web-based guided self-help for employees with depressive symptoms (Happy@Work): design of a randomized controlled trial. BMC Psychiatry 2013;13:61 [FREE Full text] [doi: 10.1186/1471-244X-13-61] [Medline: 23418886]

41. Geraedts AS, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P. Short-term effects of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014;16(5):e121 [FREE Full text] [doi: 10.2196/jmir.3185] [Medline: 24800966]

42. Happy@Work. URL: https://happy-at-work.e-behandeling.nl/ [accessed 2013-12-13] [WebCite Cache ID 6Lq0BElgt]

XSL•FO
RenderX

43. Bowman D, Scogin F, Lyrene B. The efficacy of self-examination therapy and cognitive bibliotherapy in the treatment of mild to moderate depression. Psychotherapy Research 1995 Jan;5(2):131-140. [doi: 10.1080/10503309512331331256]

44. Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive Therapy of Depression. New York: Guilford Press; 1979.

45. Franck E, Wiezer N. Handleiding Preventie Leidraad [Manual for Prevention Guideline]. Hoofddorp: TNO; 2004.

46. Franck E, Wiezer N. Ervaringen met de Preventie Leidraad [Experiences with the Prevention Guideline]. Hoofddorp: TNO; 2004.

47. Bouma J, Ranchor AV, Sanderman R, van Sonderen E. Het meten van symptomen van depressie met de CES-D; een handleiding [Assessment of Symptoms of Depression by Means of the CES-D: A Manual]. Groningen: Stichting Drukkerij Regenboog; 1995.

48. Furukawa T, Hirai T, Kitamura T, Takahashi K. Application of the Center for Epidemiologic Studies Depression Scale among first-visit psychiatric patients: a new approach to improve its performance. J Affect Disord 1997 Oct;46(1):1-13. [Medline: 9387082]

49. Haringsma R, Engels GI, Beekman AT, Spinhoven P. The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. Int J Geriatr Psychiatry 2004 Jun;19(6):558-563. [doi: 10.1002/gps.1130] [Medline: 15211536]

50. Wada K, Tanaka K, Theriault G, Satoh T, Mimura M, Miyaoka H, et al. Validity of the Center for Epidemiologic Studies Depression Scale as a screening instrument of major depressive disorder among Japanese workers. Am J Ind Med 2007 Jan;50(1):8-12. [doi: 10.1002/ajim.20403] [Medline: 17096372]

51. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory Manual. Palo Alto, CA: Consulting Psychologists Press; 1996.

52. Schaufeli WB, van Dierendonck D. UBOS: Utrechtse Burnout Schaal [UBOS: Utrecht Burnout Scale]. Lisse: Swets Test Publishers; 2000.

53. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemans B. Manual Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness (TIC-P). Rotterdam: Institute for Medical Technology Assessment; 2002.

54. van den Brink M, van den Hout WB, Stiggelbout AM, Putter H, van de Velde CJ, Kievit J. Self-reports of health-care utilization: diary or questionnaire? Int J Technol Assess Health Care 2005;21(3):298-304. [Medline: 16110708]

55. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med 2003 Feb;45(2):156-174. [Medline: 12625231]

56. Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE, et al. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. J Occup Environ Med 2004 Jun;46(6 Suppl):S23-S37. [Medline: 15194893]

57. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]

58. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med 1997 Mar;27(2):363-370. [Medline: 9089829]

59. World Health Organization. Composite International Diagnostic Interview (version 2.1). Geneva: World Health Organization; 1990.

60. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (Text Revision). Washington, DC: Jaypee Brothers Medical Publishers (P) Ltd; 2000.

61. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: L Erlbaum Associates; 1988.

62. Ensel WM. Social Support, Life Events, and Depression. New York: Academic Press; 1986:51-70.

63. Unützer J, Patrick DL, Marmon T, Simon GE, Katon WJ. Depressive symptoms and mortality in a prospective study of 2,558 older adults. Am J Geriatr Psychiatry 2002;10(5):521-530. [Medline: 12213686]

64. Zich JM, Attkisson CC, Greenfield TK. Screening for depression in primary care clinics: the CES-D and the BDI. Int J Psychiatry Med 1990;20(3):259-277. [Medline: 2265888]

65. Twisk J, de Boer M, de Vente W, Heymans M. Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. J Clin Epidemiol 2013 Sep;66(9):1022-1028. [doi: 10.1016/j.jclinepi.2013.03.017] [Medline: 23790725]

66. Phillips R, Schneider J, Molosankwe I, Leese M, Foroushani PS, Grime P, et al. Randomized controlled trial of computerized cognitive behavioural therapy for depressive symptoms: effectiveness and costs of a workplace intervention. Psychol Med 2014 Mar;44(4):741-752 [FREE Full text] [doi: 10.1017/S0033291713001323] [Medline: 23795621]

67. Warmerdam L, van Straten A, Twisk J, Riper H, Cuijpers P. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 2008;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094] [Medline: 19033149]

68. van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]

69. Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 2002 Sep;181:208-213 [FREE Full text] [Medline: 12204924]

70. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. Am Psychol 1992 Sep;47(9):1102-1114. [Medline: 1329589]

71. 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]

72. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]

73. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered Internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet 2009 Aug 22;374(9690):628-634. [doi: 10.1016/S0140-6736(09)61257-5] [Medline: 19700005]

74. 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]

75. 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]

Abbreviations

CAU: care as usual
CCBT: computer cognitive behavior therapy
CES-D: Center for Epidemiological Studies Depression scale
CT: cognitive therapy
HADS: Hospital Anxiety and Depression Scale
HPQ: Health and Work Performance Questionnaire
ITT: intention-to-treat
LMM: linear mixed modeling
MBI: Maslach Burnout Inventory-General Scale
PST: problem-solving treatment

Original Paper

A Web-Based Peer-Modeling Intervention Aimed at Lifestyle Changes in Patients With Coronary Heart Disease and Chronic Back Pain: Sequential Controlled Trial

Rebecca Schweier[1], MA; Matthias Romppel[2], Dr rer med; Cynthia Richter[1], Dipl-Soz; Eike Hoberg[3], MD; Harry Hahmann[4], MD; Inge Scherwinski[5], MD; Gregor Kosmützky[6], MD; Gesine Grande[2], Dr PH

[1]Faculty of Architecture and Social Sciences, University of Applied Sciences Leipzig (HTWK Leipzig), Leipzig, Germany

[2]Institute for Public Health and Nursing Research, Department Prevention and Health Promotion, Faculty 11: Human and Health Sciences, University of Bremen, Bremen, Germany

[3]Mühlenbergklinik Holsteinische Schweiz, Bad Malente-Gremsmühlen, Germany

[4]Klinik Schwabenland, Isny-Neutrauchburg, Germany

[5]MEDIAN Klinik Bad Lausick, Bad Lausick, Germany

[6]Reha-Zentrum Bad Driburg – Klinik Berlin, Bad Driburg, Germany

Corresponding Author:
Rebecca Schweier, MA
Faculty of Architecture and Social Sciences
University of Applied Sciences Leipzig (HTWK Leipzig)
PF 301166
Leipzig, 04251
Germany
Phone: 49 341 3076 3203
Fax: 49 341 3076 853203
Email: schweier@sug.htwk-leipzig.de

Abstract

Background: Traditional secondary prevention programs often fail to produce sustainable behavioral changes in everyday life. Peer-modeling interventions and integration of peer experiences in health education are a promising way to improve long-term effects in behavior modification. However, effects of peer support modeling on behavioral change have not been evaluated yet. Therefore, we implemented and evaluated a website featuring patient narratives about successful lifestyle changes.

Objective: Our aim is to examine the effects of using Web-based patient narratives about successful lifestyle change on improvements in physical activity and eating behavior for patients with coronary heart disease and chronic back pain 3 months after participation in a rehabilitation program.

Methods: The lebensstil-aendern ("lifestyle-change") website is a nonrestricted, no-cost, German language website that provides more than 1000 video, audio, and text clips from interviews with people with coronary heart disease and chronic back pain. To test efficacy, we conducted a sequential controlled trial and recruited patients with coronary heart disease and chronic back pain from 7 inpatient rehabilitation centers in Germany. The intervention group attended a presentation on the website; the control group did not. Physical activity and eating behavior were assessed by questionnaire during the rehabilitation program and 12 weeks later. Analyses were conducted based on an intention-to-treat and an as-treated protocol.

Results: A total of 699 patients were enrolled and 571 cases were included in the analyses (control: n=313, intervention: n=258; female: 51.1%, 292/571; age: mean 53.2, SD 8.6 years; chronic back pain: 62.5%, 357/571). Website usage in the intervention group was 46.1% (119/258). In total, 141 trial participants used the website. Independent *t* tests based on the intention-to-treat protocol only demonstrated nonsignificant trends in behavioral change related to physical activity and eating behavior. Multivariate regression analyses confirmed belonging to the intervention group was an independent predictor of self-reported improvements in physical activity regularity (β=.09, *P*=.03) and using less fat for cooking (β=.09, *P*=.04). In independent *t* tests based on the as-treated protocol, website use was associated with higher self-reported improvements in integrating physical activity into daily routine (*d*=0.22, *P*=.02), in physical activity regularity (*d*=0.23, *P*=.02), and in using less fat for cooking (*d*=0.21, *P*=.03). Multivariate regression analyses revealed that using the website at least 3 times was the only factor associated with improved lifestyle behaviors.

XSL•FO
RenderX

Conclusions: Usage of the lebensstil-aendern website corresponds to more positive lifestyle changes. However, as-treated analyses do not allow for differentiating between causal effects and selection bias. Despite these limitations, the trial indicates that more than occasional website usage is necessary to reach dose-response efficacy. Therefore, future studies should concentrate on strategies to improve adherence to Web-based interventions and to encourage more frequent usage of these programs.

Introduction

Noncommunicable and chronic diseases, such as coronary heart disease or chronic back pain, create a substantial personal and public burden globally. Diseases of the circulatory system account for 35% of deaths in Europe [1]. Chronic back pain is one of the most frequently occurring diseases of modern civilization; estimated 1-month prevalence is more than 20% globally [2]. Lifestyle-related risk factors, such as physical inactivity and unhealthy eating, are associated with a higher risk for these chronic diseases [3-7]. Thus, lifestyle change is a central aim in secondary prevention. Traditional secondary prevention programs, such as inpatient rehabilitation programs or back therapy training, have proved to be successful in the short term, but they often fail to generate sustainable behavioral changes in everyday life [8,9].

Peer-modeling interventions and integration of peer support and peer experiences in health education are a promising way to improve long-term effects in disease management [10-13]. According to social cognitive theory [14], complex behavioral patterns can be learned by watching others. In research concerning patients' expectations and information-seeking behavior, patients often report a special interest in patient narratives and peer support [15-17]. Peer support is shown to be effective in improving self-efficacy and recovery from surgery in patients with heart diseases [18,19] as well as reducing disability and pain in chronic pain patients [20]. Interventions featuring patient narratives have a positive effect on cancer screening decisions [21] and self-care in diabetes patients [22,23].

Patient narratives about illness experiences can be found on various Web forums, such as video and social networking websites (eg, YouTube and Facebook), and on publicly funded websites such as the healthtalkonline website (formerly DIPEx) [24]. However, empirical data on the impact of patient narratives and peer modeling on lifestyle behavior modification are still lacking. Therefore, we implemented and subsequently evaluated a website featuring patient narratives about successful lifestyle changes. Our objective was to examine whether the use of this website [25] contributes to improvements in physical activity and eating behavior in chronically ill coronary heart disease and back pain patients.

Methods

The Website

Overview

The website lebensstil-aendern ("lifestyle-change") is a nonrestricted, no-cost, German language website that provides more than 1000 video, audio, and text clips from interviews with patients living with coronary heart disease and chronic back pain. The website has been online since November 2011 and is certified with the Health on the Net Foundation Code of Conduct (HONcode), a certificate addressing reliability and trustworthiness of medical information on the Internet [26]. Details on implementation, usability evaluation, and usage statistics are reported elsewhere [27].

The Patient Narratives

We interviewed 39 people with coronary heart disease and 27 with chronic back pain who reported that they had successfully modified their behavior in at least 1 lifestyle domain for more than 6 months. These problem-centered interviews [28] focused on strategies and barriers in maintaining a healthy lifestyle. Interviewees were recruited across Germany through media coverage, information events, flyers, word-of-mouth recommendations, and social media.

The interviews lasted between 1 to 3 hours. We extracted short clips addressing different aspects of lifestyle modification. All clips are provided with text and may also contain a video or audio clip depending on the interviewees' preferences. Before publication, patients were contacted again to decide which clips should be published, whether statements should be removed, and if they wanted to use a pseudonym or their real name. For quality assessment and to avoid potentially harmful suggestions, experienced cardiologists and orthopedists reviewed all patient statements before publication.

Website Structure and Content

The website is structured using a horizontal menu with links to the home page, a news page, the patient narratives, a forum, background information, and a contact form (see Figure 1).

The news page is updated several times a month by the project team and provides news about recent research results and announcements of new patient narratives, recipes, and project-related updates. The patient narratives are divided into 2 indication-specific modules and structured using a vertical menu with the following categories: overcoming your "weaker self", getting active, eating healthier, reducing stress, getting

support, dealing with the disease, quitting smoking (only in the coronary heart disease module), and keeping the spine in mind (only in the chronic back pain module). In addition to the menu, suitable clips can be found via a filter for age and sex, a tag cloud, searching for keywords, or through overview pages for

each patient. Users can comment on single clips and evaluate them (see Figure 2).

The forum contains the same categories as the patient narratives. Posts are accessible for everyone, but users must be registered and logged in to post to the forum.

Figure 1. Screenshot of the lebensstil-aendern home page.

Figure 2. Screenshot of a video subpage from the coronary heart disease module. On the left side is the vertical menu with the tag cloud beneath. Below the video player are buttons for evaluation and short information text about the patient's health condition.
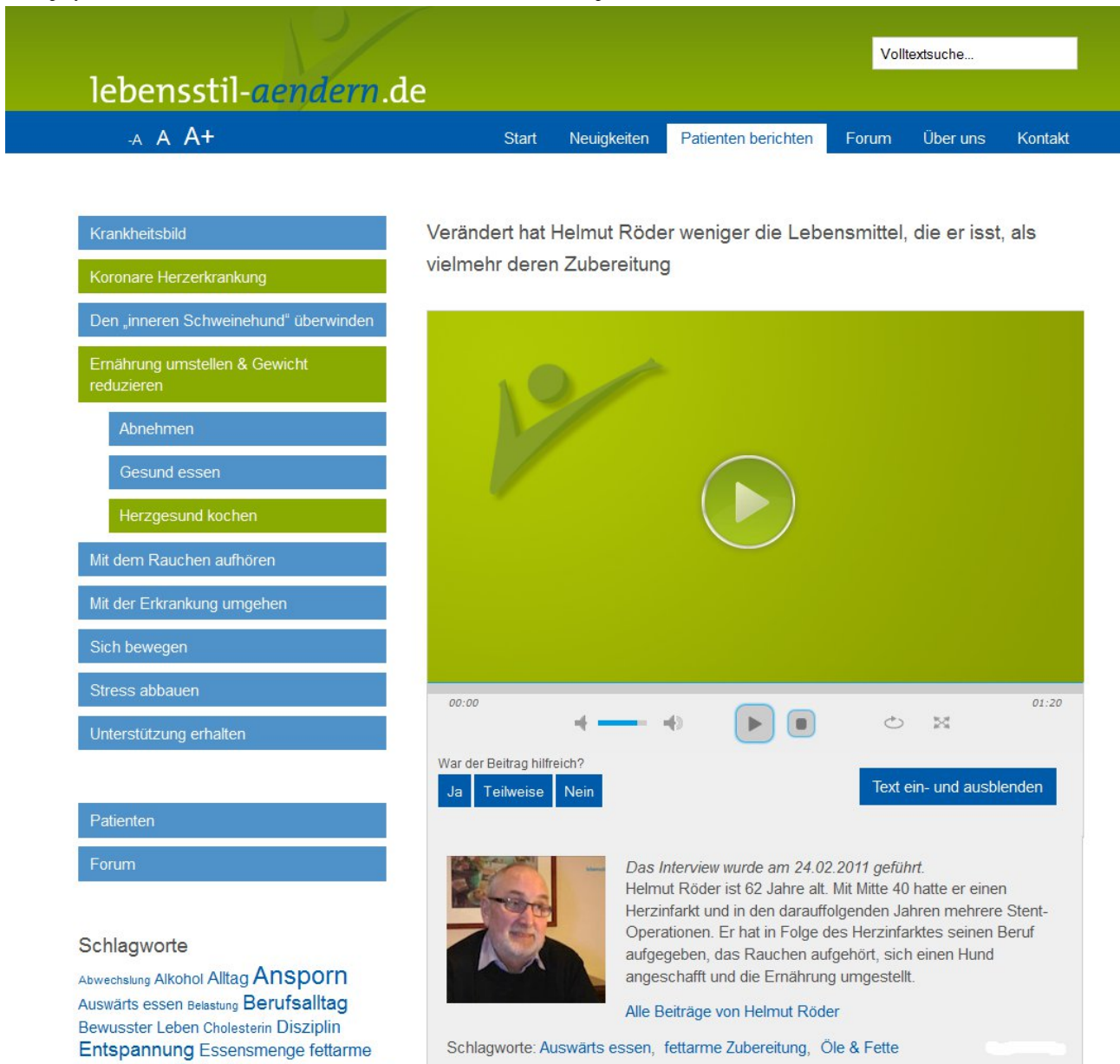
Recruitment

We recruited patients in 7 inpatient rehabilitation centers from September 2012 to August 2013. Patients had to be diagnosed with coronary heart disease (International Statistical Classification of Diseases and Related Health Problems, ICD-10 I20-25) or chronic back pain (ICD-10 M50-54), have sufficient German language skills, and have no disabling cognitive deficits.

All patients meeting the inclusion criteria were enrolled and sequentially assigned to the control group and intervention group, respectively (see Figure 3). Sequential allocation was chosen to avoid contamination between groups. The study protocol was approved by the Ethics Committee of the Faculty of Medicine, Leipzig University (reference number: 301-10-04102010).

Figure 3. Flowchart of study participants.
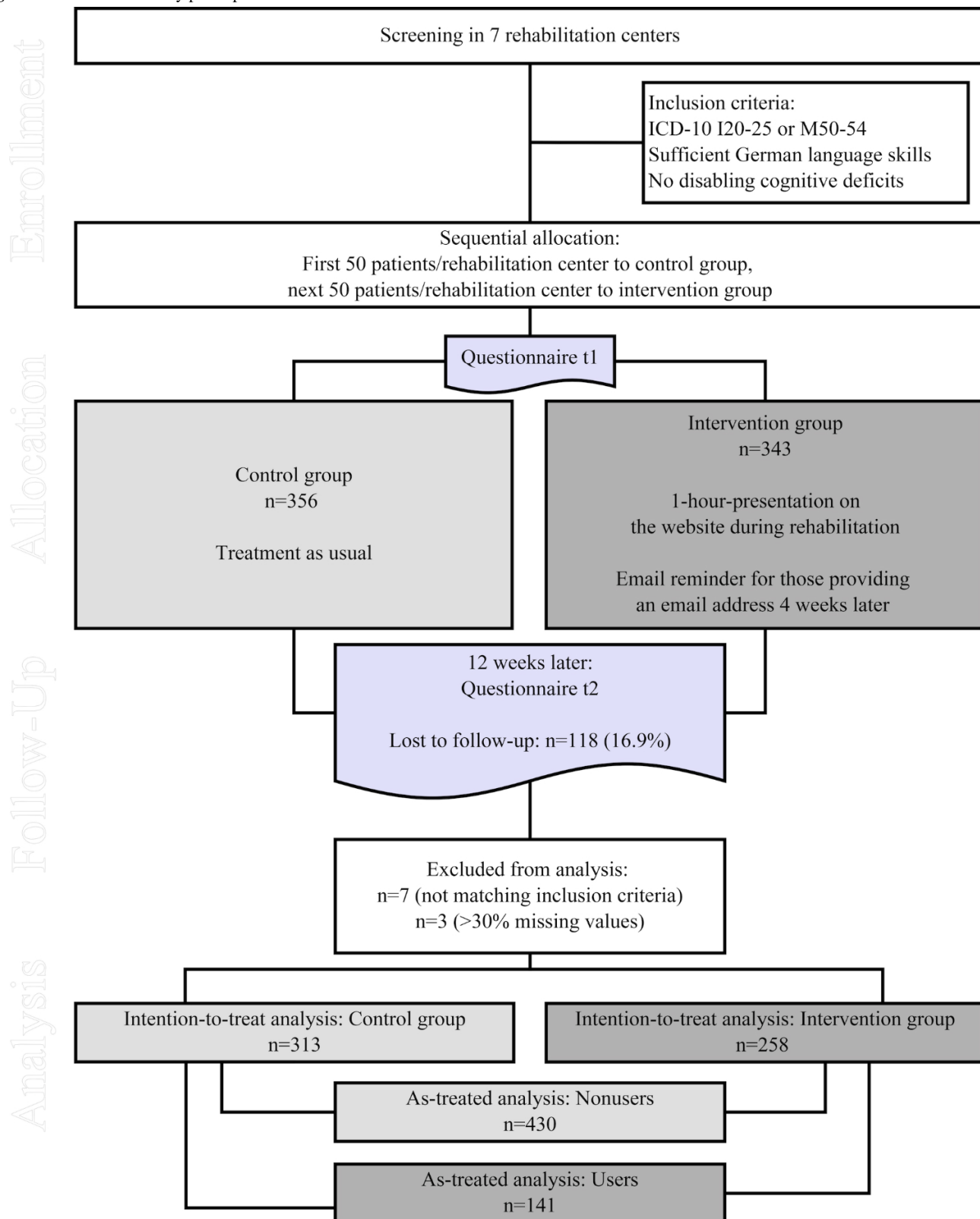
Intervention

Patients in the intervention group were invited to join a presentation about the website. During the 1-hour presentation, the project team introduced the aims and scopes of the project and provided information on how to find suitable content on the website. They also discussed how to register and post on the forum, and addressed issues of data protection, including anonymity. To give the participants an idea of the website content, 1 to 3 patient narrative videos were shown. To further reduce barriers to using the website, the intervention group received a detailed printed manual and was encouraged to contact the project team in case of problems or questions. Those participants who provided an email address received an email reminder with the link to the website 4 weeks after the

presentation. Patients in the control group received no intervention.

Measures

Both the intervention and control groups completed a questionnaire at some point during participation in a rehabilitation program. Questions addressed sociodemographic characteristics (eg, age, sex, education, and income), diagnosis, body mass index (BMI), and baseline behavior for physical activity and eating routine. The latter 2 were measured on a numerical scale of 0 to 10 asking about the frequency of exercise and the attention paid to a healthy diet. The follow-up questionnaire also included 2 questions about exercise frequency and attention paid to healthy diet. Additionally, patients were asked to rate, on a 5-point Likert scale ranging from slightly deteriorated (–1) to improved substantially (3), their improvements in physical activity (frequency, regularity, integration in daily routine) and eating behavior (eating more healthy foods, avoiding unhealthy foods, eating smaller portions, using less fat for cooking). Lastly, usage of and satisfaction with the website were assessed.

Statistical Analysis

Statistical analysis was conducted using SPSS Statistics version 21 (IBM Corp, Armonk, NY, USA). Uncertainties during data entry (eg, the participants' handwriting was difficult to decipher or more than 1 box marked as an answer) were discussed by the team and resolved by consensus. Seven participants were excluded for not matching the ICD-10 diagnoses I20-25 or M50-54 and all 118 dropout cases and 3 other cases were excluded for having more than 30% missing data. Missing data in the remaining sample were estimated using multiple imputations under a fully conditional specification and 10 iterations. Because 41.3% (236/571) cases had missing values in any of the variables to be imputed, we imputed 60 datasets [29]. For evaluation, we compared groups by using independent t tests and chi-square tests and calculated Cohen's d, P values, and 95% confidence intervals for the mean difference between groups (Δ). For regression analyses, regression weights (B) plus their 95% confidence intervals and beta values (standardized regression weights) are reported.

For estimating improvements in exercise frequency and attention paid to a healthy diet, we calculated mean differences (Δ) between the 2 measure points and compared these mean differences in independent t tests. Results are reported both for an intention-to-treat analysis (intervention group vs control group) as well as for an as-treated analysis (website users vs website nonusers). For the regression analyses on the various outcomes in the intention-to-treat analysis, we included intervention group, indication of coronary heart disease, male sex, age per decade, more than 10 years of education, income per €200, BMI, and baseline behavior (exercise frequency for physical activity items and attention toward a healthy diet for eating behavior items). For the regression analyses on the various outcomes in the as-treated analysis, we replaced the intervention group variable with frequent website usage (at least 3 times) and occasional website usage (once or twice).

Results

Dropout Analysis

A total of 699 participants filled out the first questionnaire during rehabilitation. During follow-up, 118 patients (16.8%) dropped out. Dropout was higher among patients with coronary heart disease (21.5% vs 14.1%, χ^2_1=6.3, P=.01) and in the intervention group (23.2% vs 11.1%, χ^2_1=18.1, P<.001). On average, those who dropped out were 1.9 years younger (t_{685}=2.11, P=.04) and had a €189 lower income (t_{623}=2.74, P=.006).

Website Usage

In total, 24 of 313 control group participants (7.7%) and 164 of 258 intervention group participants (63.6%) claimed to know the website. Nevertheless, approximately one-quarter of these patients (46/188, 24.5%) never visited the website. Overall, 119 of 258 in the intervention group (46.1%) and 22 of 313 in the control group (7.0%) visited the website; 1 participant (0.7%) stated he used the website 3 to 4 times a week, 6 participants (4.3%) stated they used it once or twice a week, 60 (42.6%) used it once or twice a month, and 74 (52.5%) accessed the website at least once.

Intention-to-Treat Analysis

Participants had a mean age of 53.2 years (SD 8.6), 51.1% (292/571) were women, and 62.5% (357/571) had been diagnosed with chronic back pain (Table 1). Intervention group participants had higher levels of education than control group participants. Income was distributed equally in both samples, but it was slightly lower than the average German net equivalent income of €1835 per month in 2012 [30]. Participants had a preobese BMI of 28.1 kg/m^2 (SD 5.2), much higher than the German average (25.7 kg/m^2 in 2009 [31]).

There were no significant differences in outcome variables between the groups (see Table 2). The intervention group showed a tendency to improve the regularity of their physical activity more than the control group but this did not meet statistical significance in t tests. In the regression analysis, belonging to the intervention group emerged as an independent predictor for more recognizable improvements in regularity of physical activity, including younger age, higher income, lower baseline exercise frequency, and prevalence of coronary heart disease compared to chronic back pain (see Table 3). In addition to this, we found a trend in t tests in favor of the intervention group for using less fat for cooking. In the regression analysis, belonging to the intervention group also emerged as an independent predictor, accompanied by independent effects of prevalence of coronary heart disease compared to chronic back pain, lower education, and higher BMI.

XSL•FO
RenderX

Table 1. Baseline characteristics—intention-to-treat analysis.

Variable	Control group (n=313)	Intervention group (n=258)	Total (n=571)	t_{569}	χ^2_1	P
Age (years), mean (SD)	52.7 (8.1)	53.8 (9.2)	53.2 (8.6)	−1.52		.13
Sex, n (%)						
Female	170 (54.3)	122 (47.3)	292 (51.1)			
Male	143 (45.7)	136 (52.7)	279 (48.9)		0.2	.69
Indication, n (%)						
Coronary heart disease	115 (36.7)	99 (38.4)	214 (37.5)			
Chronic back pain	198 (63.3)	159 (61.6)	357 (62.5)		2.8	.10
Education, n (%)						
≤10 years of school	236 (75.5)	172 (66.8)	408 (71.5)			
>10 years of school	77 (24.5)	86 (33.2)	163 (28.5)		4.9	.03
Net equivalent income per month (€), mean (SD)	1679 (621)	1713 (633)	1695 (627)	−0.62		.56
BMI (kg/m^2), mean (SD)	28.3 (5.3)	27.9 (5.2)	28.1 (5.2)	0.82		.41
Exercise frequency (scale 0-10), mean (SD)	3.89 (2.88)	4.27 (2.81)	4.06 (2.85)	−1.61		.11
Attention paid to a healthy diet (scale 0-10), mean (SD)	5.61 (2.37)	5.59 (2.60)	5.60 (2.47)	0.10		.92

Table 2. Independent t tests—intention-to-treat analysis.

Dependent variable	Control group (n=313)	Intervention group (n=258)	Mean difference	t_{569}	95% CI	Cohen's d
Physical activity						
Frequency of doing exercise, Δ (t2–t1)	0.77	0.51	−0.26	3.17	−0.79, 0.27	−0.08
Improvements in physical activity frequency, mean	1.26	1.33	0.07	0.92	−0.09, 0.23	0.08
Improvements in physical activity regularity, mean	1.19	1.33	0.14	0.98	−0.03, 0.30	0.14
Improvements in physical activity in daily routine, mean	1.23	1.31	0.08	0.97	−0.08, 0.24	0.08
Eating behavior						
Attention paid to a healthy diet, Δ (t2–t1)	1.10	1.20	0.11	2.34	−0.29, 0.51	0.04
Improvements in eating more healthy foods , mean	1.31	1.32	0.01	0.88	−0.14, 0.17	0.01
Improvements in eating less unhealthy foods, mean	1.32	1.39	0.08	0.95	−0.09, 0.24	0.08
Improvements in eating smaller portions, mean	1.08	1.10	0.03	0.89	−0.13, 0.18	0.03
Improvements in using less fat for cooking, mean	1.17	1.30	0.13	0.88	−0.03, 0.30	0.14

Table 3. Multivariate linear regression—intention-to-treat analysis. Each line starting with a dependent variable contains information from one regression analysis (N=571) and provides model fit (R^2), results for the variable group (regardless of significance), and those variables which were demonstrated to be significant independent predictors in the model.

Dependent variable	R^2	Group B (95% CI)	β	Variable	B (95% CI)	β
Physical activity						
Frequency of doing exercise, Δ (t2–t1)	.42	0.01 (–0.40, 0.42)	.001	Indication	0.63 (0.12, 1.13)	.10
				Age	–0.40 (–0.65, –0.14)	–.11
				Baseline behavior	–0.70 (–0.78, –0.63)	–.63
Improvements in physical activity frequency	.06	0.10 (–0.06, 0.26)	.05			
				Age	–0.18 (–0.28, –0.09)	–.17
				Income	0.04 (0.01, 0.07)	.13
				Baseline behavior	–0.03 (–0.06, 0.00)	–.09
Improvements in physical activity regularity	.06	0.18 (0.02, 0.34)	.09	Indication	0.21 (0.01, 0.42)	.11
				Age	–0.18 (–0.28, –0.08)	–.16
				Baseline behavior	–0.03 (–0.06, 0.00)	–.10
Improvements in physical activity in daily routine	.06	0.11 (–0.05, 0.27)	.08	Age	–0.13 (–0.23, –0.03)	–.12
				Education	–0.28 (–0.46, –0.09)	–.13
Eating behavior						
Attention paid to a healthy diet, Δ (t2–t1)	.44	0.12 (–0.19, 0.42)	.02	Baseline behavior	–0.63 (–0.70, –0.57)	–.65
Improvements in eating more healthy foods	.08	0.03 (–0.12, 0.18}	.02	BMI	0.04 (0.02, 0.05)	.20
Improvements in eating less un-healthy foods	.09	0.08 (–0.08, 0.23)	.04	Indication	0.29 (0.09, 0.48)	.15
				BMI	0.03 (0.02, 0.05)	.18
Improvements in eating smaller portions	.09	0.05 (–0.10, 0.20)	.03	BMI	0.05 (0.04, 0.07)	.29
Improvements in using less fat for cooking	.07	0.16 (0.01, 0.31)	.09	Indication	0.20 (0.01, 0.39)	0.11
				Education	–0.22 (–0.40, –0.05)	–.11
				BMI	0.03 (0.02, 0.04)	.19

As-Treated Analysis

Users had higher levels of education than nonusers. We did not find any significant differences in baseline variables between nonusers and users in the as-treated analysis (see Table 4).

XSL·FO
RenderX

Table 4. Baseline characteristics—as-treated analysis.

Variable	Nonusers (n=430)	Users (n=141)	t_{569}	χ^2_1	P
Age (years), mean (SD)	53.5 (8.6)	52.2 (8.4)	−1.56		.12
Sex, n (%)					
Female	217 (50.5)	75 (53.2)			
Male	213 (49.5)	66 (46.8)		0.3	.57
Indication, n (%)					
Coronary heart disease	165 (38.4)	49 (34.8)			
Chronic back pain	265 (61.6)	92 (65.2)		0.6	.44
Education, n (%)					
≤10 years of school	317 (73.7)	92 (65.2)			
>10 years of school	113 (26.3)	49 (34.8)		3.9	.049
Net equivalent income per month (€), mean (SD)	1667 (626)	1780 (622)	−1.81		.07
BMI (kg/m^2), mean (SD)	27.9 (5.1)	28.8 (5.5)	1.72		.09
Frequency of doing exercise, (scale 0-10), mean (SD)	4.09 (2.89)	3.98 (2.74)	−0.38		.71
Attention paid to a healthy diet (scale 0-10), mean (SD)	5.59 (2.45)	5.62 (2.54)	0.14		.89

Website use was associated with larger increases in physical activity: users were more successful in integrating physical activity into their daily routine and in improving their physical activity regularity. We did not find a significantly higher mean difference in the frequency of doing exercise between baseline and follow-up 12 weeks later.

Website use was not associated with higher mean differences in the attention paid to a healthy diet, but it was with more success in dietary behaviors (ie, using less fat for cooking) (Table 5).

Table 5. Independent *t* tests—as-treated analysis.

Dependent variable	Nonusers (n=430)	Users (n=141)	Mean difference	t_{569}	95% CI	Cohen's *d*
Physical activity						
Frequency of doing exercise, Δ (t2–t1)	0.60	0.81	0.21	0.67	−0.40, 0.82	0.07
Improvements in physical activity frequency, mean	1.26	1.40	0.14	1.56	−0.04, 0.32	0.15
Improvements in physical activity regularity, mean	1.20	1.42	0.22	2.33	0.04, 0.41	0.23
Improvements in physical activity in daily routine, mean	1.21	1.43	0.21	2.24	0.03, 0.40	0.22
Eating behavior						
Attention paid to a healthy diet, Δ (t2–t1)	1.09	1.32	0.23	0.99	−0.23, 0.69	0.10
Improvements in eating more healthy foods, mean	1.28	1.43	0.15	1.64	−0.03, 0.33	0.16
Improvements in eating less unhealthy foods, mean	1.31	1.47	0.16	1.65	−0.03, 0.34	0.16
Improvements in eating smaller portions, mean	1.04	1.22	0.17	1.91	−0.05, 0.35	0.19
Improvements in using less fat for cooking, mean	1.18	1.38	0.20	2.17	0.02, 0.38	0.21

In the multivariate regression analyses (Table 6), we found small independent effects of frequent website use on improvements in physical activity regularity and improvements in integration of physical activity into daily routine. We also found small effects of frequent website use on all 5 variables measuring eating behavior—frequent website use was an independent predictor for increased attention paid to a healthy diet, improved use of healthy and avoidance of unhealthy foods, improvements in eating smaller portions, and reduced use of fat for cooking.

XSL•FO
RenderX

Table 6. Multivariate linear regression—as-treated analysis. Each line starting with a dependent variable contains information from one regression analysis (N=571) and provides model fit (R^2), results for the variables occasional website usage and frequent website usage (regardless of significance), and those variables which were demonstrated to be significant independent predictors in the model.

Dependent variable	R^2	Occasional website usage		Frequent website usage		Further independent predictors		
		B (95% CI)	β	B (95% CI)	β	Variable	B (95% CI)	β
Physical activity								
Frequency of doing exercise, Δ (t2–t1)	.42	–0.13 (–0.74, 0.49)	–.01	0.28 (–0.36, 0.92)	.03	Indication	0.63 (0.13, 1.14)	.10
						Age	–0.41 (–0.66, –0.15)	–.11
						Baseline behavior	–0.70 (–0.77, –0.62)	–.63
Improvements in physical activity frequency	.06	–0.005 (–0.23, 0.22)	–.002	0.21 (–0.03, 0.45)	.07	Age	–0.18 (0.28, 0.09)	–.17
						Income	0.04 (0.01, 0.07)	.13
Improvements in physical activity regularity	.06	0.12 (–0.12, 0.36)	.04	0.27 (0.02, 0.51)	.09	Indication	0.21 (0.01, 0.41)	.11
						Age	–0.18 (–0.28, –0.08)	–.16
						Baseline behavior	–0.03 (–0.06, 0.00)	–.09
Improvements in physical activity in daily routine	.07	0.06 (–0.17, 0.30)	.02	0.39 (0.14, 0.63)	.13	Age	–0.13 (–0.23, –0.03)	–.12
						Education	–0.27 (–0.46, –0.09)	–.13
Eating behavior								
Attention paid, a healthy diet, Δ (t2–t1)	.43	–0.14 (–0.59, 0.31)	–.02	0.73 (0.26, 1.21)	.10	Baseline behavior	–0.64 (–0.71, –0.58)	–.66
Improvements in eating more healthy foods	.09	–0.06 (–0.28, 0.17)	–.02	0.36 (0.12, 0.59)	.12	BMI	0.03 (0.02, 0.05)	.19
						Baseline behavior	–0.04 (–0.07, 0.00)	–.10
Improvements in eating less unhealthy foods	.10	0.04 (–0.20, 0.27)	.01	0.32 (0.08, 0.56)	.11	Indication	0.29 (0.10, 0.48)	.15
						BMI	0.03 (0.02, 0.05)	.17
Improvements in eating smaller portions	.10	0.03 (–0.20, 0.25)	.01	0.28 (0.05, 0.51)	.10	BMI	0.05 (0.03, 0.06)	.28
Improvements in using less fat for cooking	.08	0.02 (–0.21, 0.25)	.01	0.40 (0.17, 0.64)	.14	Indication	0.20 (0.01, 0.39)	.11
						Education	–0.21 (–0.39, –0.03)	–.10
						BMI	0.03 (0.02, 0.05)	.18

Discussion

Principal Results

The aim of this study was to examine the effects of a Web-based intervention featuring patient narratives about successful lifestyle changes on physical activity and eating behavior for chronically ill coronary heart disease and back pain patients. We conducted a sequential controlled trial in which the intervention group participated in a presentation about the website during rehabilitation.

In the intention-to-treat analysis, there were no significant effects of the intervention in the bivariate analyses and only small interventional effects on physical activity regularity and on using less fat for cooking in the multivariate analyses. However, all but 1 (frequency of doing exercise) of the measured outcome variables showed positive tendencies in the expected direction.

Comparing website users to nonusers in the as-treated analysis, we found an association between website usage and improvements in most physical activity and dietary behavior outcome variables. Multivariate regression analyses revealed that this association persists only in conjunction with frequent website usage (defined as having visited the website at least 3 times) not with occasional website usage.

Implications and Future Research

A larger sample size, associated with a sufficiently large power, may have ensured the detection of even small intervention effects. In any case, effect sizes were consistently small and demand a closer look at possible reasons. First, the small intervention effects in the intention-to-treat analysis can be explained in part by the low usage of the website because less than half of the intervention group visited the website. The problem of low usage rates is well known and an often cited problem in trials on Web-based interventions [32-35]. Some personal factors, such as young age, male gender, low income, and low levels of education, are associated with lower rates of preventive service utilization in general [36], further increasing the socially determined health-related disadvantages. In our study, only a lower educational level was related to lower rates of website usage. Thus, other groups who are often difficult to reach seem to benefit from the website.

In addition to the low utilization rates, the intervention dose was probably too low to result in behavioral modifications for most participants. The results of the as-treated analysis indicate that more than occasional website usage is necessary to reach dose-response efficacy. This finding is consistent with results from other studies [37-39] that also found a positive relation between the number of log-ins and outcomes. In these studies, the necessary number of log-ins also was not reached among all users.

Finally, although the integration of peer experiences is seen as promising in health education and disease management and meets the interest of patients with a chronic condition, the additional inclusion of further behavior modification strategies (eg, individual tailoring or providing feedback) may be required in a peer-modeling approach to show satisfactorily large effects on lifestyle behavior [40].

Future studies should concentrate on strategies to improve adherence to Web-based interventions and especially to induce more frequent usage of these programs. Individually tailored interventions may be a promising approach because these interventions, besides being more efficacious, have also been shown to stimulate more frequent use of Web-based interventions and to increase adherence [21,41].

Limitations

We were able to assess behavioral change through self-reported nonvalidated measures only, which are subject to response bias. One important aim of rehabilitation programs is to increase awareness for unhealthy lifestyle behavior. Thus, recruiting the participants at this point of time might have led to an overestimation of the 2 baseline variables, which then would have a flattening effect on the mean difference between t2 and t1. Due to time restrictions in the research project, we could only observe the effects of the intervention 3 months after participation in the rehabilitation program. The long-term effects of the intervention remain unclear. Furthermore, we did not control for Internet literacy or Internet usage patterns, which might explain differences in website usage rates and website efficacy.

The as-treated analysis implicates positive results, but does not allow for a causal interpretation because nonmeasured variables associated with more frequent use of the website could account for the improved health behavior or successful patients might tend to more actively seek support and information and use a website such as the lebensstil-aendern website.

For the multivariate regression analyses, model fit was very low. This could either be due to variables accounting for behavioral change but not being included in the model or, considering that all outcome variables were single item variables, unreliable measurements contributing to the large error variances.

Conclusions

To our knowledge, our study is the first trial to test the efficacy of Web-based patient narratives on behavioral outcomes in coronary heart disease and chronic back pain patients. In our study, patients that use the website more frequently report more favorable health-related behavior and more marked improvements in physical activity and diet. Even if more motivated patients seek support and information more actively, for this patient group in particular, a website such as the lebensstil-aendern website appears to be a helpful tool.

Authors' Contributions

Rebecca Schweier wrote the manuscript. Cynthia Richter and Rebecca Schweier implemented the website and planned and conducted the trial together. Matthias Romppel helped in planning the trial and performed the statistical analysis with Rebecca Schweier. Eike Hoberg, Harry Hahmann, Inge Scherwinski, Gregor Kosmützky, and their teams did the screening in the rehabilitation centers, recruited the control group, and helped in conducting the trial. Gesine Grande managed the research project and supervised during each of its respective stages.

XSL•FO

RenderX

Conflicts of Interest

None declared.

References

1. World Health Organization. European Health Report 2012: Charting the Way to Well-Being. Copenhagen: World Health Organization Regional Office for Europe; 2013. URL: http://www.euro.who.int/__data/assets/pdf_file/0003/184161/The-European-Health-Report-2012,-FULL-REPORT-w-cover.pdf [WebCite Cache ID 6RGEzMOkO]

2. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. Arthritis Rheum 2012 Jun;64(6):2028-2037 [FREE Full text] [doi: 10.1002/art.34347] [Medline: 22231424]

3. De Bacquer D, Dallongeville J, Kotseva K, Cooney MT, Pajak A, Deckers JW, et al. Residual risk of cardiovascular mortality in patients with coronary heart disease: the EUROASPIRE risk categories. Int J Cardiol 2013 Sep 30;168(2):910-914. [doi: 10.1016/j.ijcard.2012.10.051] [Medline: 23157810]

4. Carlsson AC, Wändell PE, Gigante B, Leander K, Hellenius ML, de Faire U. Seven modifiable lifestyle factors predict reduced risk for ischemic cardiovascular disease and all-cause mortality regardless of body mass index: a cohort study. Int J Cardiol 2013 Sep 30;168(2):946-952. [doi: 10.1016/j.ijcard.2012.10.045] [Medline: 23181992]

5. Chow CK, Jolly S, Rao-Melacini P, Fox KA, Anand SS, Yusuf S. Association of diet, exercise, and smoking modification with risk of early cardiovascular events after acute coronary syndromes. Circulation 2010 Feb 16;121(6):750-758 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.891523] [Medline: 20124123]

6. Smuck M, Kao MC, Brar N, Martinez-Ith A, Choi J, Tomkins-Lane CC. Does physical activity influence the relationship between low back pain and obesity? Spine J 2014 Feb 1;14(2):209-216. [doi: 10.1016/j.spinee.2013.11.010] [Medline: 24239800]

7. Neuhauser H, Ellert U, Ziese T. Chronische Rückenschmerzen in der Allgemeinbevölkerung in Deutschland 2002/2003: Prävalenz und besonders betroffene Bevölkerungsgruppen [Chronic Back Pain in the General Population in Germany 2002/2003: Prevalence and Highly Affected Population Groups]. Gesundheitswesen 2005;67(10):685-693. [doi: 10.1055/s-2005-858701]

8. Kotseva K, Wood D, De Backer G, De Bacquer D, Pyörälä K, Keil U, EUROASPIRE Study Group. EUROASPIRE III: a survey on the lifestyle, risk factors and use of cardioprotective drug therapies in coronary patients from 22 European countries. Eur J Cardiovasc Prev Rehabil 2009 Apr;16(2):121-137. [doi: 10.1097/HJR.0b013e3283294b1d] [Medline: 19287307]

9. van Tulder MW, Koes B, Malmivaara A. Outcome of non-invasive treatment modalities on back pain: an evidence-based review. Eur Spine J 2006 Jan;15 Suppl 1:S64-S81 [FREE Full text] [doi: 10.1007/s00586-005-1048-6] [Medline: 16320031]

10. Turner G, Shepherd J. A method in search of a theory: peer education and health promotion. Health Educ Res 1999 Apr;14(2):235-247 [FREE Full text] [Medline: 10387503]

11. Gallant MP. The influence of social support on chronic illness self-management: a review and directions for research. Health Educ Behav 2003 Apr;30(2):170-195. [Medline: 12693522]

12. Heisler M. Different models to mobilize peer support to improve diabetes self-management and clinical outcomes: evidence, logistics, evaluation considerations and needs for future research. Fam Pract 2010 Jun;27 Suppl 1:i23-i32 [FREE Full text] [doi: 10.1093/fampra/cmp003] [Medline: 19293400]

13. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [Medline: 15222150]

14. Bandura A. Sozial-kognitive Lerntheorie. 1st ed. Stuttgart: Klett-Cotta; 1979.

15. Sillence E, Briggs P, Herxheimer A. Personal experiences matter: what patients think about hypertension information online. He@lth Information on the Internet 2004;42(1):3-5 [FREE Full text]

16. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [doi: 10.1016/S0738-3991(03)00116-2] [Medline: 15062905]

17. Newman MA, Ziebland S, Barker KL. Patients' views of a multimedia resource featuring experiences of rheumatoid arthritis: pilot evaluation of www.healthtalkonline.org. Health Informatics J 2009 Jun;15(2):147-159. [doi: 10.1177/1460458209102974] [Medline: 19474226]

18. Colella TJ, King KM. Peer support. An under-recognized resource in cardiac recovery. Eur J Cardiovasc Nurs 2004 Sep;3(3):211-217. [doi: 10.1016/j.ejcnurse.2004.04.001] [Medline: 15350230]

19. Parry M, Watt-Watson J. Peer support intervention trials for individuals with heart disease: a systematic review. Eur J Cardiovasc Nurs 2010 Mar;9(1):57-67. [doi: 10.1016/j.ejcnurse.2009.10.002] [Medline: 19926339]

20. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. Pain 2011 Aug;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012] [Medline: 21565446]

21. Jensen JD, King AJ, Carcioppolo N, Krakow M, Samadder NJ, Morgan S. Comparing tailored and narrative worksite interventions at increasing colonoscopy adherence in adults 50-75: a randomized controlled trial. Soc Sci Med 2014 Mar;104:31-40. [doi: 10.1016/j.socscimed.2013.12.003] [Medline: 24581059]

22. Brownson CA, Heisler M. The role of peer support in diabetes care and self-management. Patient 2009 Mar 1;2(1):5-17. [doi: 10.2165/01312067-200902010-00002] [Medline: 22273055]

23. Campbell T, Dunt D, Fitzgerald JL, Gordon I. The impact of patient narratives on self-efficacy and self-care in Australians with type 2 diabetes: stage 1 results of a randomized trial. Health Promot Int 2013 Aug 28. [doi: 10.1093/heapro/dat058] [Medline: 23985247]

24. Ryan K, Herxheimer A. Personal Experiences of Health and Illness on the Web: A Resource for Patients, Carers and Health Professionals. Journal of the Malta College of Pharmacy Practice 2007(13):27-30.

25. lebensstil-aendern.de [lifestyle-change.de]. URL: http://lebensstil-aendern.de/ [accessed 2014-07-17] [WebCite Cache ID 6R8OQutou]

26. Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites. Comput Biol Med 1998 Sep;28(5):603-610. [Medline: 9861515]

27. Richter C, Schweier R, Romppel M, Grande G. Erfolgreiche Peers als Modell. Ein internetbasiertes Unterstützungsangebot zur nachhaltigen Lebensstilmodifikation. Phys Rehab Kur Med 2013 Oct 15;23(05):283-291. [doi: 10.1055/s-0033-1355347]

28. Witzel A. The problem-centered interview. Forum Qualitative Sozialforschung [Forum: Qualitative Social Research] 2000;1(1):22 [FREE Full text]

29. Bodner TE. What improves with increased missing data imputations? Structural Equation Modeling: A Multidisciplinary Journal 2008 Oct 22;15(4):651-675. [doi: 10.1080/10705510802339072]

30. Statistisches Bundesamt. Leben in Europa (EU-SILC): Einkommen und Lebensbedingungen in Deutschland und der Europäischen Union [Living in Europe (EU-SILC) – Income and Living Conditions in Germany and the European Union 2012]. 2014. URL: https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/LebeninEuropa/EinkommenLebensbedingungen2150300127004.pdf?__blob=publicationFile [accessed 2014-05-22] [WebCite Cache ID 6PkrM14Ee]

31. Federal Statistical Office. 2009. Behavioural and Risk Aspects of Health: Height, Weight and Body Mass Index of the Population by Sex and Age-Groups – Results of Microcensus 2009 URL: https://www.destatis.de/EN/FactsFigures/SocietyState/Health/HealthStatusBehaviourRelevantHealth/Tables/BodyMassIndex.html [accessed 2014-03-26] [WebCite Cache ID 6OMVmMAs6]

32. 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]

33. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]

34. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. Patient Prefer Adherence 2008;2:57-65 [FREE Full text] [Medline: 19920945]

35. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]

36. Dryden R, Williams B, McCowan C, Themessl-Huber M. What do we know about who does and does not attend general health checks? Findings from a narrative scoping review. BMC Public Health 2012;12:723 [FREE Full text] [doi: 10.1186/1471-2458-12-723] [Medline: 22938046]

37. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]

38. Lewis B, Williams D, Dunsiger S, Sciamanna C, Whiteley J, Napolitano M, et al. User attitudes towards physical activity websites in a randomized controlled trial. Prev Med 2008 Nov;47(5):508-513 [FREE Full text] [doi: 10.1016/j.ypmed.2008.07.020] [Medline: 18773915]

39. Schulz DN, Kremers SP, Vandelanotte C, van Adrichem MJ, Schneider F, Candel MJ, et al. Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes. J Med Internet Res 2014;16(1):e26 [FREE Full text] [doi: 10.2196/jmir.3094] [Medline: 24472854]

40. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]

41. Aalbers T, Baars MA, Rikkert MG. Characteristics of effective Internet-mediated interventions to change lifestyle in people aged 50 and older: a systematic review. Ageing Res Rev 2011 Sep;10(4):487-497. [doi: 10.1016/j.arr.2011.05.001] [Medline: 21628005]

Abbreviations

BMI: Body mass index
ICD: International Statistical Classification of Diseases and Related Health Problems

XSL•FO
RenderX

Viewpoint

Purple: A Modular System for Developing and Deploying Behavioral Intervention Technologies

Stephen M Schueller[1], PhD; Mark Begale[1], BA; Frank J Penedo[2], PhD; David C Mohr[1], PhD

[1]Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Northwestern University, Chicago, IL, United States
[2]Northwestern University, Department of Medical Social Sciences, Chicago, IL, United States

Corresponding Author:
David C Mohr, PhD
Center for Behavioral Intervention Technologies
Department of Preventive Medicine
Northwestern University
680 N Lake Shore Drive
Suite 1400
Chicago, IL, 60611
United States
Phone: 1 312 503 1403
Fax: 1 312 503 0982
Email: d-mohr@northwestern.edu

Abstract

The creation, deployment, and evaluation of Web-based and mobile-based applications for health, mental health, and wellness within research settings has tended to be siloed, with each research group developing their own systems and features. This has led to technological features and products that are not sharable across research teams, thereby limiting collaboration, reducing the speed of dissemination, and raising the bar for entry into this area of research. This paper provides an overview of Purple, an extensible, modular, and repurposable system created for the development of Web-based and mobile-based applications for health behavior change. Purple contains features required to construct applications and to manage and evaluate research trials using these applications. Core functionality of Purple includes elements that support user management, content authorship, content delivery, and data management. We discuss the history and development of the Purple system guided by the rationale of producing a system that allows greater collaboration and understanding across research teams interested in investigating similar questions and using similar methods. Purple provides a useful tool to meet the needs of stakeholders involved in the creation, provision, and usage of eHealth and mHealth applications. Housed in a non-profit, academic institution, Purple also offers the potential to facilitate the diffusion of knowledge across the research community and improve our capacity to deliver useful and usable applications that support the behavior change of end users.

KEYWORDS

Introduction

Health care stakeholders are increasingly turning to information and communication technologies to support the development and deployment of interventions aimed at supporting behavior change related to health, mental health, and wellness. These interventions, which we refer to as "behavioral intervention technologies" (BITs) [1,2], make use of Internet-connected devices, such as desktop computers, mobile devices, and/or sensors to provide low-cost, scalable, and effective methods to expand the portfolio of tools available and to enhance the reach, impact, and speedy implementation of current research and practice [3,4]. While the literature on the clinical application of BITs has been growing at an enormous rate, much of the discussion on development methods has been confined to the engineering and computer science communities through conference proceedings of meetings such as CHI (Computer-Human Interaction), UbiComp (Ubiquitous Computing), Pervasive Health, and MobiSys (Mobile Systems). BIT development involves considerable challenges [2], not the least of which stems from the necessary collaboration between

the clinical researchers, who most often develop the idea for the intervention, and technologists who subsequently create the BIT. Despite the growing number of research teams and resultant BITs, most efforts in this space lack a generalizable technological approach to their development, often producing divergent platforms and software to meet similar needs, solve similar problems, and create BITs with similar sets of features. This paper, intended for a broad readership, describes Purple, a modular system that supports the development and deployment of BITs.

Purple is an iteratively constructed system that has evolved to address the needs of stakeholders who create, deploy, and use BITs, including researchers, technologists, clinicians, and patients. We begin with an overview of Purple's underlying rationale and history. Because the needs of stakeholders guided the development of Purple's capabilities, we present a brief description of BIT stakeholders as they relate to Purple. We then provide a brief review of Purple's current technical features. To give the reader a more concrete understanding of these features, we present two case studies of BITs developed using Purple. The diversity of these case studies in clinical approach and technological componentry illustrates the flexibility of Purple. Yet the use of generalizable components illustrates the usefulness of taking a modular approach with an eye towards repurposing components. This paper aims to aid stakeholders interested in this space by outlining the necessary components of a BIT development system and presenting the features of Purple that seek to address each of these needs.

Purple

Overview

The field of digital health generally, and BIT research specifically, is an emerging area of transdisciplinary research that requires specific expertise from multiple disciplines as well as the development of shared conceptual frameworks that transcend the perspectives of any one field [5]. Purple aims to address these needs by providing a framework that supports the development of BITs in a manner that is, as much as possible, extensible, modular, and repurposable for a community of researchers. This framework allows individual stakeholders to make use of the contributions and development provided by others and contribute based on their specific expertise.

Many BITs, even when targeting very different behaviors and clinical outcomes, require similar features. These features include methods of enrolling users, user authentication, and content delivery, and the ability to receive, store, and manipulate user information and feed it back to users and investigators on a variety of devices (eg, mobile phones, computers, and tablets). Currently, individual research labs that develop BITs as one-off projects often recreate these features repeatedly. Although some labs have the capacity to build modular systems and repurpose code for subsequent projects, this is the exception rather than the rule with regards to development related to clinical research. Thus, one goal of Purple was to build a system that leverages resources across multiple researchers, labs, and institutions. Indeed, other resources exist that target some of the same goals of Purple, including Open mHealth [6], LifeGuide [7], and the Behavior Wizard [7]. These include acting as a repository of developed code for mHealth projects [6], providing conceptual guidance [8], and reducing the need for programming knowledge via open-source software that allows the creation and modification of BITs using authoring tools and logic commands [7]. Purple arose concurrently with these resources. However, as none provided a sufficient solution for the problems encountered by clinical researchers (eg, eliminating the need for a programmer) or sufficiently covered the user management, recruitment, and data acquisition needs for our various projects, we developed Purple to facilitate the development of our own projects as well as external ones.

We first created generalizable features that addressed the required elements for any BIT (eg, user management, recruitment, and data acquisition) and the capabilities to support the development of new features, which over time can be formed into generalizable features, and then be made available to researchers. Developing each of these features could take anywhere between a few days to a few weeks of developer time depending on whether a similar feature (ie, with existing available code) developed for another need (eg, a news site or blog editor) could be sufficiently modified or if the required feature needs to be constructed from scratch. Through sharing and iteration of features and the knowledge associated with developing, deploying, and evaluating them, Purple provides avenues for the dissemination of technical and design information outside of traditional academic papers and conferences. Table 1 displays a list of currently available technological features.

Table 1. Technical features of Purple.

Features	Creators	End users	Essential	Optional
User management				
User authentication				
Study IDs, passwords	X	X	X	
Password generation	X	X	X	
User definition				
Personal health information management	X	X	X	
Study management				
Study/intervention assignment, randomization	X		X	
Content authorship/management				
Didactic content				
Content editor	X			X
Generic form builder	X			X
Content type (eg, text, images, video, audio)	X	X		X
Progression (eg, sequential, non-sequential)	X		X	
Assessments				
Self-report types (checkbox, slider, text, etc)		X		X
Phone sensor acquisition		X		X
External data API access		X		X
Intra-assessment logic (sequential, adaptive, etc)	X		X	
Scoring (summing, complex algorithms)	X	X	X	
BIT construction				
Progression rule definition (eg, chronological, task-based, time-based, custom)	X	X	X	
Custom interaction tool authorship	X			X
Content delivery				
Platform (phone, tablet, Web, etc)	X	X	X	
Connectivity (online, offline, both)	X	X	X	
Interfaces (touch, mouse, keyboard, etc)	X	X	X	
Data management				
Access	X		X	
Visualization (charting, custom, InfoVis)	X	X		X

The use of Purple involves processes that have been and continue to be developed to facilitate collaboration between clinical investigators and technologists across the life of a project [5,9,10]. Because BIT development commonly uses iterative development processes, new features are developed with the maximum degree of flexibility possible to allow investigators to adapt the characteristics of those features over the development process. The need for the Purple architecture was born out of our experience. Like many investigators, we began work using a commercial developer who built a website that met our specifications for an Internet intervention for depression [11]. Through evaluation, we encountered many flaws in our design and conceptualization, we developed new ideas regarding methods of engaging users, and we found colleagues who wanted to build on our work. Unfortunately, the one-off construction we used limited our ability to adapt, improve, and share our work. Our next project, an intervention for cancer survivors called Onward [12], employed a development strategy that prioritized modularity and flexibility. The features developed within Onward became the beginning of the feature set available in Purple. Features from Onward are now used to further development on Onward-like interventions and have been adapted for several Web-based and tablet interventions targeting a variety of clinical problems, such as PC CHIP, a tablet-based intervention described below. Similarly, our work in mobile interventions began with the retooling of Mobilyze, an intervention for depression built for Nokia's Symbian operating system [13] for the Android platform. This was done concurrently with work on FOCUS, a smartphone intervention for schizophrenia [14]. The work completed on these two

projects created an infrastructure of mobile features for Purple that is now being used by multiple projects. Initial efforts in the creation of Web and mobile features developed within the modular Purple framework were given a significant boost in 2011, with the establishment of the Center for Behavioral Intervention Technologies (CBITs), with support from the National Institute of Mental Health and additional funding from Northwestern University.

With the establishment of CBITs, we began to expand our services, and by extension the community of researchers contributing to and deriving benefit from Purple, to investigators outside of Northwestern University. As of this writing, CBITs supports more than 50 projects in the United States, Africa, and South America, funded by more than 20 National Institutes of Health (NIH) grants, 2 grants from the Canadian government, and numerous other government agencies, foundations, and universities. The expansion of research teams using and contributing to Purple helps achieve its overarching vision to become a hub for the community of BIT researchers that allows for the transfer of development methods via recycling of code across research labs, allowing the researchers to rapidly benefit from the knowledge gained by the larger community of researchers, as well as the larger group of stakeholders.

Another aspirational vision for Purple is to develop databases that will allow the community of researchers to answer questions that no individual lab can address in isolation. For example, lab-based research allows investigators to examine how specific features or characteristics are related to adherence and outcomes within their own intervention. These data can be examined across research using meta-analytic techniques (eg, [15]); however, as meta-analyses combine only aggregated data, they have limited power to detect effects and can only address questions at a gross level, using the kinds of information that are collected routinely across research times and reported in publications. In cases where individual subject data are available, the data can be pooled to conduct meta-analyses on the original data, also known as mega-analysis [16]. Mega-analyses are preferable to meta-analyses when only a few relevant studies exist or subgroup analyses are desired. Both of these are often true in the domain of BITs. The speed of technological innovation means researchers must base decisions on conclusions drawn from few, and often outdated, studies. Furthermore, researchers are often interested in how the user groups they are targeting (eg, clinical needs, population, setting) might use and benefit from particular BITs or BIT elements. The most powerful mega-analyses would come from studies that share a common framework that guides methodological decisions (eg, delivery of treatment components, outcomes assessed). Within the Purple system, use and outcome data are consistently defined and application features and characteristics are consistently identified. While users of Purple are not required to share data, if they choose to do so, the data are added to a de-identified database. This database provides the opportunity to evaluate the relationship between specific design elements, as well as use and clinical outcomes. As this database grows, it will provide a resource to enhance the evidence base that could inform how best to design BITs for general clinical

purposes, and for specific populations and settings. As it stands, several collaborators have stored their data on our servers; however, the proper ontologies still need to be created to make use of this data. As such, data have yet to be combined and analyzed cross-projects, and this remains an aspiration rather than an application.

Purple developers have made an intentional decision to maintain this system in a non-profit university setting. The primary goal is to support research rather than to develop a product. As research requires testing intervention principles, Purple is optimized for the extreme data needs of researchers, including the collection of detailed data to permit fine-grained analyses of use patterns of study participants. Furthermore, research focuses on the creation and dissemination of knowledge. As developed components are available to be independently disseminated, they are opensourced to the community at large. For example, "Purple Robot", a real-time mobile phone sensor data acquisition platform that can collect all available data from Android phones, is currently available for download on the GooglePlay store with source code available at GitHub [17].

Stakeholders Involved in Behavioral Intervention Technologies

The creation of a development framework must consider the needs of all stakeholders, who can be involved in the design, development, funding, deployment and use of BITs. We present in Figure 1 the highest level of classes—creators, end users, and purveyors—as overlapping categories, with more specific stakeholders selected who may play multiple roles within the process of developing and deploying BITs, represented within these categories [18]. The creator class encompasses all stakeholders who construct, shape, improve, and disseminate a behavioral intervention technology and can be grouped into subclasses such as clinical researchers, technologists, and funders. End users make use of the content and resources provided by BIT creators. They are the ones who actually engage with the intervention in order to facilitate the behavior changes desired by the creators. End users may include patients or public consumers, clinical providers who interface with the applications, or peers who are networked through the application. End users also author the data necessary to evaluate the efficacy of the BIT at reaching its stated goal. This includes both use data that correspond to an end user's engagement with the BIT and clinical data that correspond to the specific clinical and behavior change targets. Purveyors are those responsible for the dissemination the intervention and may include app stores or care providing organizations, as well as policy makers. Many stakeholders may have roles in several classes, which is why the highest categories are presented as overlapping within Figure 1. For example, clinical care organizations are commonly purveyors but may also be involved in the creation and the use of interventions.

Although features of Purple have been developed with most of these stakeholders in mind, for the purposes of the current paper we focus most directly on creators, who use Purple to instantiate their intentions and aims into code and produce a BIT, and end users whose behavior the BIT is intended to support or change.

Figure 1. Stakeholders involved in BITs.
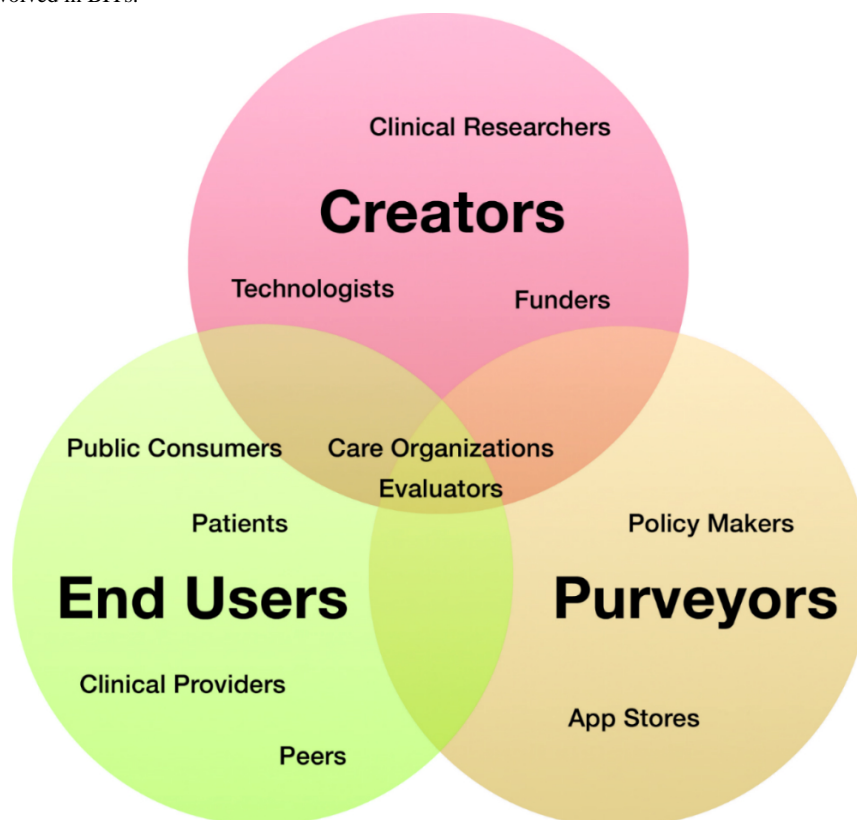
Technical Features

Overview

Over time, Purple has expanded to include a range of technical features. Table 1 displays many of these features and groups them into four categories representing core functionality: user management, content authorship, content delivery, and data management. The table includes an indication of the class of stakeholders each feature is aimed at and whether that feature is essential or optional to the functioning of a BIT in the Purple System. "Essential" features are those that are necessary to deploy and evaluate a BIT in a research context and thus appear in every BIT developed using Purple (eg, user management, recruitment/enrollment). These features have been developed to be generalizable across projects because of their recurring need in each. Optional features are unique to a given project and created in response to specifications of the producers for that BIT. These features are typically developed as non-generalizable features that over time might be able to be modified for use across projects. Last, Purple contains many generalizable but optional features. Generalizable but optional features are created when the demand across projects is such that it makes sense to create a generalizable solution like a dashboard to allow a clinician or coach to view data and manage BIT users (eg, [19]) or through the development of an initially non-generalizable feature into a generalizable one. This table provides a menu of options of the current capacities of Purple, but new capacities are frequently developed in response to new problem spaces or through new collaborations. We discuss each of these feature groups in more detail below.

User Management

User management is an important element for any BIT to ensure that data are structured and interpretable to allow for evaluation. As such, the BIT system must track who enters the system, when, and be able to provide access to the correct information and content for each individual user and stakeholder. This is also crucial for practice to ensure privacy and security concerns, promote trust in the BIT [20,21], and support user experience by ensuring that it works properly when content is tailored or personalized.

To support privacy and security, an essential feature within Purple is to allow users to create unique identifiers, passwords, and to assign these users to a specific study. In Purple, these processes make up user authentication features. Another aspect of user management is defining variables related to that user and specifying how these variables correspond to other features within the intervention. For example, Purple user management includes the ability to define which stakeholders are permitted access to demographic or personal health information (PHI). Some instances of PHI stored within Purple should not be viewable by most personnel and stakeholders. For instance, email addresses or phone numbers of end users may be required for Purple to deliver automated email or text messages; however, most people engaging with or receiving data from Purple do not require access to such information and therefore should not see it. User authentication also involves validating the user's credential against different features of the intervention including the date the intervention started, condition membership (eg, experimental vs control condition), and other user specific permissions. This allows the intervention provided to each user to differ according to conditions preset by the creators.

Content Authorship

Content authorship tools enable creators with limited technological sophistication to participate more directly and easily in the development process. This is useful for creating BITs, as experts in behavior change often do not have the capability to program BITs on their own and need to team with technologists who do. We use the term content broadly to refer to anything that is provided to a user (eg, text, video, or audio) for didactic purposes; for example, text "lessons" that are commonly used in BITs such as Internet sites or mobile applications [13,22], assessments, notifications (email, secure messaging, in-phone notifications), or feedback. Although many behavioral scientists might see these as functionally different, we discuss them together because, from the vantage of the Purple system, they use similar or identical technical features. For example, the primary difference between a piece of didactic content and an assessment within Purple is that the didactic content only displays text, images, video and/or audio, while assessment combines this with another element that allows the user to input information.

Purple contains several content authorship tools to meet the needs of the creator. These can be generally classified as tools to create didactic content, assessments, and BIT construction. As much as possible, these tools are designed so that individuals without programming knowledge can create basic didactic content, notifications, or modify elements of interactive tools to meet the more specific requirements of the application. The content editor allows the creator to enter text and upload images, video, and audio, which will then be displayed back to the end user. The editor is useful at creating BITs that draw on specific types of scripted interactions (eg, a user progressing through a series of slides). As a majority of BITs use similar types of interactions, content authorship tools provide a useful low-effort way to script and pilot content and intervention ideas before progressing to more advanced programming required in more complicated and BIT-specific interaction flows. We use examples below to demonstrate how content authorship tools are used in two different BITs, but across projects these tools are ones that empower clinical researchers to create and iterate aspects of the intervention.

Assessment refers to the acquisition and scoring of data. The content editor also allows the creator to input the content of questions for self-report data and to determine the presentation of those questions, such as the number of questions on a page and the type of response required, such as a check box, slider, or free text. A variety of intra-assessment logic types are available, including simple logic such as sequential or random presentation of assessment items, adaptive presentation making use of Boolean logic such as if-then or more complex logic, or item-response theory. Purple also has the capacity to pipe responses from one question into the text of another question.

Purple allows for the collection of a wide variety of additional data important for use and usability evaluation, as well as ubiquitous computing. Within Purple, detail log and use data are collected from sensors and applications running on a website or mobile device, as well as data from other sources via application programming interfaces (APIs). Purple Robot [17]

is a mobile application within the stack of Purple applications that supports the acquisition of a full range of mobile phone sensors (eg, global positioning system [GPS], Wi-Fi, accelerometers, Bluetooth, altimeters, light sensors), device information (eg, battery, running software, and apps), and options to scan for visible external devices (eg, Bluetooth devices or wireless access points). Purple Robot also provides a JavaScript API that has been used to link external devices (eg, Fitbit, Jawbone, wireless pillboxes) and allows for additional sensors to be added in the future. In order to facilitate the transfer of data into meaningful information, Purple offers a variety of scoring options. These range from simple summation to more complex algorithms and machine learning functions.

These fundamental tools for the creation of didactic content, assessment, and logic can be combined to create interactive tools. Interactive tools developed so far using Purple include goal-setting tools, activity scheduling and monitoring tools (using calendars), and tools that promote thought restructuring. These elements can also be modified and combined in new ways to create custom interaction tools. For example, altering a tool from tracking food to tracking activity can alter an intervention from targeting diet for weight loss to behavioral activation for depression. A generic form builder within Purple creates many of these tools. The generic form builder allows a creator to provide free text boxes or other modes of data entry and then to provide this information back to user at specified times. The generic form builder is also able to repurpose code for different modes of entry (such as on a calendar or on a feed) that have been developed for various projects within Purple.

Progression rules refer to the rules that define how a user moves through an intervention. While some BITs allow access to all elements of an intervention from the first login, others contain rules that sequence the presentation of treatment elements. While complex progression rules can require programming effort, the editor also contains features that allow creators to set basic progression rules. For example, the creator may specify that lesson, tool, or notification should be available to the end user after performing specific tasks, reaching or not reaching a treatment goal, or simply the passage of a specific amount of time. Creators can also define within each lesson when specific pages should be available to end users. This includes sequential logic definitions (eg, page 2 displays after page 1) or non-sequential logic (page 3 displays only after a certain condition is met, such as number of views or score on an assessment). These tools can also be applied to tailoring interventions to user characteristics.

Content Delivery

In the context of BITs, treatment elements can be delivered in multiple modalities to multiple users (eg, patients/consumers, coaches, clinicians) and according to previously specified progression rules (as discussed above). Some intervention delivery methods have been requested and used in so many of the BITs designed within Purple that features have been developed specifically for those purposes. These include calendar displays and graphing tools. Other content varies significantly from BIT to BIT but requires the researcher to make specific decisions to shape the end user's experience.

Content delivery essentially enables the intervention to be played to the end user on their device or computer based on rules defined in the content authorship and management elements in Purple. Content delivery in the Purple system uses HTML5-based hybrid applications. Hybrid applications allow for the production of a current generation, Web-based experience for end users without writing specific Android or iOS code. For user-facing portions, these applications can be deployed to Android, iOS, or Windows mobile devices. This allows for deployment on multiple platforms. Designing for multiple platforms (or optimizing for a specific platform) involves several decisions that require an understanding of the desired user interaction. Some of our applications are completely on the client side (eg, they make use of the browser side and do not require an active Internet connection), whereas others are server side, thus requiring an active connection. In order to reuse the code developed for online applications, we use Adobe PhoneGap/Apache Cordova software to package the application so they work offline, allowing creators to deliver interventions that do not necessarily require connectivity. The use of HTML-5 hybrid development also supports the more rapid production of prototypes, facilitates testing directly on desktop or laptop computers, and uses a more generally available programming skillset.

In addition, by using HTML5-based hybrid applications, the user interface can be adapted to the delivery platform. For example, a traditional check box is easy to select with mouse-driven interfaces. The same box would be encased in a button for touch screen devices to simplify user selection. Thus, multiple versions of tools are created, allowing the user experience to be tailored for mobile phones, tablets, and desktop computers. Important human-computer interaction issues here might include making specific interface-related decisions, for example, using a library to make Web applications more responsive on the phone or determining what the concept of a "hover" action might be when it does not exist on a tablet but does on a desktop computer. Purple offers a system where many of these design possibilities are mapped out, allowing the designer to make decisions based on the intervention concept and needs of the target population.

Data Management

Data are used for a variety of purposes, including the evaluation of the BIT, as well as to manage and define the user experience. In general, two types of outcomes are of interest to researchers evaluating BITs: clinical outcomes and use (or adherence) outcomes. Clinical outcomes represent the treatment target, such as weight, smoking cessation, or depression. These outcomes typically use participant self-reports, or observer reports (eg, clinicians) collected via assessments conducted within the system but may also use project specific APIs or sensors. Use data can be measured in a variety of ways including the frequency that a participant has logged in or launched a BIT, the features of the BIT used, and the length of time interacting with the BIT. While Purple has the capacity to collect fine grained data of every interaction, it is usually preferable to pre-specify the use data desired in advance because collection of all available data leads to significant demands on the server and provides so much data that it is difficult to sort through, understand, analyze, and draw conclusions. Data can be made available to researchers in raw data dumps in commonly used formats such as Excel, or through direct access to databases to facilitate real-time data analysis.

End users, and sometimes researchers, often require data to be visualized to facilitate understanding. A variety of charting libraries are available that can be used to create traditional bar, line, and pie charts. Purple also allows for the development of custom authored visualizations that allow creators to develop creative methods of displaying data for specific use cases. InfoVis combines charting or custom visualizations with didactic content to create visual representations that can be effectively deployed to enhance understanding.

Purple contains sets of functionality that map onto common use cases. Novel configurations will require programmer effort until these become sufficiently standardized to allow the creation of modular elements. Within and between the BIT elements, there is a "handshake" between features that requires technological developer support. As much as possible, the editor allows non-programmer creators to develop and import content, manage assessments, define progression rules, and access data, maximizing control and management of the development process in the hands of clinical researchers. In almost all cases, we use structured query language (SQL)–backed data storage, which most modern developers can use to interface with the data directly. Some projects using Purple house data on our servers, whereas other collaborators store data on their own server.

Case Studies

Overview

The modular structure of the Purple system allows common technical features within Purple to be integrated to create unique BITs. We present here two examples of BITs developed using the Purple system, which have different clinical aims, target populations, and use different devices, to illustrate how similar Purple features can be used to create very different interventions. Table 2 displays brief descriptions of manifestations of the technical features in Purple across these two interventions. A more detailed description of each intervention follows.

Table 2. Technical features within intervention case studies.

Features	Mobilyze	PC CHIP
User management		
User authentication		
Study IDs, passwords	Performed by creator at randomization	Performed by creator at start of project
Password generation	None	Single password set for all users
User definition		
Personal health information management	Email (within Purple Robot)	Age, gender, email (for contact)
Study management		
Study/intervention assignment, randomization	Content released daily	Content unlocked weekly, questions released 1 hour after scheduled session, user specific permissions and group memberships
Content authorship/management		
Didactic content		
Content editor	Used to enter and schedule release of text lessons	Used to enter text, video, and audio
Generic form builder	Calendar tool, active coping tool, coping card	External tools used through assessment center API
Content type (eg, text, images, video, audio)	Text	Text, video, audio
Progression (eg, sequential, non-sequential)	Sequential; new lessons available daily	Sequential; new content unlocks weekly
Assessments		
Self-report types (checkbox, slider, text, etc)	All form types	Radio buttons only
Phone sensor acquisition	Various through Purple Robot	None
External data API access	None	WebEX (video conferencing)
Intra-assessment logic (sequential, adaptive, etc)	Basic (lists of questions) and advanced (scoring controls progression)	Computerized adaptive testing
Scoring (summing, complex algorithms)	Scoring occurs based on machine learning rules	Scoring uses Item-Response Theory (IRT) algorithms from Patient-Reported-Outcomes Measurement Information System (PROMIS)
BIT construction		
Progression rule definition (eg, chronological, task-based, time-based, custom)	Time-based	Time-based, linked to video conferencing session
Custom interaction tool authorship	Form-based tools can be authored using "tool builder"	Not applicable
Content delivery		
Platform (phone, tablet, Web, etc)	Smartphone	Tablet
Connectivity (online, offline, both)	Both	Both
Interfaces (touch, mouse, keyboard, etc)	Touch	Touch
Data management		
Access	Access to data is provided about sensors using a server side database, user tool data are stored separate from sensor data	User tool usage data are accessible on the server, survey data available independently through PROMIS
Visualization (charting, custom, InfoVis)	Graphs, list reviews, bubble visualizations	None

XSL•FO

RenderX

Mobilyze: A Smartphone-Based Intervention for Individuals With Depression

Mobilyze is an Android smartphone intervention based on the principles of behavior activation, which aims to reduce depressive symptoms by increasing the user's engagement in activities that are pleasurable or provide a sense of accomplishment. The ultimate goal of the Mobilyze project is to develop a context-sensing system that harnesses data from sensors embedded within the smartphone to identify user states that may be relevant to treatment, such as location, activity, social context, and mood. As such, Mobilyze consists of three major components: (1) a patient-facing application that includes lessons and application specific tools, (2) a smartphone application that collects sensor data, monitors authored models, and provides phone interactions that support context-sensing functions, and (3) a server side framework to collect unstructured and structured user data. A pilot version of Mobilyze was evaluated with 8 end users with major depressive disorder and provided promising support for this intervention model and the development of the context-sensing system [13]. Our current work on this project involves retooling the intervention for modern platforms (ie, Android, iOS) and improving functionality related to context-sensing capabilities. This updated version of Mobilyze is a 12-week intervention consisting of content and tools consistent with principles of behavioral activation for depression [23]. Mobilyze tools include daily lessons, activity tracker, mood check-in, active coping tool, coping cards, and progress graphs.

When Mobilyze is installed on a user's smartphone device, the end user does not have to enter any passwords or login credentials to open the application. The application does, however, need to be able to identify the user such that data entered will be linked to that user. This is accomplished using the email address associated with the phone collected by the Purple Robot application. A user can also go into Purple Robot and create a unique identifier that is different from their email address. For each user, creators are able to specify the date the intervention started for the user and release content daily corresponding to the number of days the user has been enrolled in the study.

Mobilyze consists of approximately 60 brief lessons, deployed every 1-2 days, that take no more than 5 minutes to read. These lessons were entered using the content editor authoring tool. Each day, lesson content is displayed as text via a widget set up on the home screen of the mobile device, and some lesson content links directly to other tools contained within the Mobilyze program. Lessons follow sequential release rules set within the editor. Each lesson is released on a specific day and then available to end users to view again throughout participation in the 12-week program. Other tools within Mobilyze make use of generalizable features developed for across-project use created using a generic form builder. These features include a calendaring tool (the Activity Tracker) and generic form-building tools (Active Coping and Coping Card supporting positive activities and management of stressful events). The activity tracker supports activity scheduling and monitoring that is part of behavioral activation, by allowing end users to report the date and time for the event and to rate how

much pleasure and accomplishment they experienced during that event. Each element of this mode of entry corresponds to things present in other aspects of the Purple system (eg, pleasure and accomplishment are rated using slide bars that could be changed via the content authoring tool; date and time are entered using a date/time entry tool). The generic form builder allows for creation of structured interactions with the program. For example the coping card consists of some instructional content (entered via the content authorship tool), a text entry box, and assigning a date and time when that coping card will be sent to the end user as a notification.

Assessments within Mobilyze include both user-entered data and sensor data collected passively by Purple Robot. For user-entered data, a single question mood check-in requires users to report their mood using a 10-point slider bar that ranges from "bad" to "good". All aspects of this check-in could be easily altered using the content authoring tools ranging from the question prompt, scale anchors, or type of entry mode (eg, slider, checkboxes, radio button, free text entry). Delivery of the check-in is controlled via the content authoring tool. In different phases of the deployment, it has been provided using time-based or context-based rules. The context sensing portion of this project aims to use sensor data collected from the phone to identify treatment-relevant user states. Raw data from the phone's sensors (eg, GPS, call logs, accelerometer, application usage, screen state) are automatically collected with no effort required by the user. This context sensing requires the ability to download raw data from phone sensors as well as other external sources such as weather data, the ability to generate models that can use data to predict user states, and to upload those models to the phone where they can be used to detect user states in real time. These features were developed into an application, called Purple Robot [17]. Purple Robot provides a JavaScript API that collects the sensor data available via Android smartphone devices (eg, GPS, accelerometer, visible Wi-Fi/Bluetooth). The sensor data are linked with labeled data drawn from the user-entered responses to the mood prompts for user information about their mood to train models.

On the server side, the Purple system collects content authored by the end user within the application as data (eg, mood scores from the check-in, events, and associated ratings from the activity tracker) and detailed data related to each user's interactions with the application. Data acquired through the sensors and use data are stored in separate databases as an additional precaution of data security and confidentiality. Use data includes when, what, and how a user clicks on each button within the Mobilyze application. For example, Purple tracks when users click on a lesson to open it, when they click on each button to progress to the next page within the lesson, and when they click on a tool or home screen button at the end of the lesson to move to another section of the Mobilyze program. From this, the creator knows when and how long a user is reading each lesson or interacting with each tool. This information is fed back to members of the study team. Coaches have an interface that provides use and clinical data that support brief calls focused on maintaining motivation and discussing aspects of the individual end user's use of the application based on data entered into the program. Values entered via the various

tools are fed back to the user as graphs displaying things such as mood over time or average pleasure and accomplishment in activities over time.

For Mobilyze, the existence of Purple Robot and the content management system allowed some development of features without additional programming but with additional content generation by the clinical team (eg, writing lessons, writing question prompts, providing feedback on visual interface). Other aspects, for example, the homepage of the application, the timing of assessment prompts, mapping servers to facilitate data transfer and storage, require programming. As such, Purple speeds the development of Mobilyze and provides a framework so that similar Mobilyze-like interventions can be more easily developed in the future.

PC CHIP: A Tablet-Based Intervention for Prostate Cancer Survivors

PC CHIP is a tablet-based intervention that aims to improve health-related quality of life and treatment-related symptom burden in men diagnosed with advanced or metastatic prostate cancer. The intervention consists of cognitive-behavioral stress management techniques that incorporate interactive didactics, role-plays, and content and emotional processing exercises. The intervention components also include relaxation training exercises, stress awareness, coping and communication skills provision, and social support with an emphasis on treatment-related side effects and interpersonal disruption [24]. PC CHIP is currently being evaluated in a randomized clinical trial comparing the tablet-based stress management intervention to a health promotion attention matched control condition. Throughout the 10-week intervention period, users receive weekly content through a mix of live video sessions led by a group facilitator, text slides, and audio recordings in a group format with 4-6 participants. In the health promotion condition, participants receive only didactics on health promotion topics such as the benefits of nutrition, physical activity, and treatment compliance, and other salient topics that offer information to the participants but do not involve any stress management or cognitive and emotional processing.

As users are enrolled, they are assigned to groups. Study IDs and passwords are generated by the creator at the start of the each of the group's enrollment. Group passwords and logins are used rather than individual passwords and logins because content is consistent across end users in each group. Thus, an end user's study ID and passwords serve to identify their group. Content is released weekly and also related to group-specific permissions. Questions related to the video sessions, for example, are released 1 hour after the scheduled session for that group.

PC CHIP makes use of several content authorship and management features within Purple. Purple's content editor (displayed in Figure 2) allows entry of video content and PowerPoint slides that are then synched to each other (see Figure 3). In this way, users are able to receive content in multiple modalities at the same time. The editor is also used to upload audio files that contain the relaxation exercises and other didactics summarizing key aspects of the materials presented in the live video conferencing sessions. Progression rules within

PC CHIP include both sequential and non-sequential processes. Throughout the 10-week trial, new content is available for review each week. A daily counter associates the entire intervention to a countdown clock based on a weekly video conferencing session delivered to the end users. As a result of this clock, weekly assessments and lesson content are made available directly following a weekly session.

Assessments are seamlessly provided using an external service that uses the Patient Reported Outcomes Measurement Information System (PROMIS) instruments and the Assessment Center toolkit [25]. The PROMIS Assessment Center toolkit is a free online data collection tool that allows researchers to create study-specific websites to make use of various features of the PROMIS instruments including customization of items, real-time scoring of computer adaptive tests. PROMIS instruments undergo updates and minor modifications from the PROMIS research team thus allowing researchers who use the toolkit access to the most up-to-date measures without the need to make updates themselves. Thus, connecting PC CHIP to this toolkit through Purple provides a considerable benefit with initial integration as the only programming consideration. In the PC CHIP trial, all participants are asked to log in and complete PROMIS-based computerized adaptive testing for symptoms of fatigue, pain, depression, anxiety, and physical functioning. Purple technology also provides system-generated alerts to participants with reminders to complete the PROMIS instruments. Additional API access is used to link with the WebEX video conferencing platform. WebEX is a Cisco product that supports the various needs of the PC CHIP project (ie, a single group facilitator with multiple participants). At the end of the 10 weeks, the users receive an automatically generated post-assessment evaluation that taps into the participants' overall comfort in the sessions and perceived efficacy and understanding of the session materials. These assessments consist of radio buttons created using the Purple content authoring system.

For the PC CHIP project, Purple collects data related to use of the various features (eg, accessing WebEX, viewing presentations, accessing assessments) and the results of the post-assessment evaluation. These data are accessible on the server and provided to the creators for review of how the application is functioning, which aspects might be more beneficial, and users' perceptions of the intervention. Outcome data are available independently through the PROMIS Assessment Center toolkit [25]. As such, no visualizations of outcome data are provided to the end users.

In the case of PC CHIP, a majority of the intervention is configured in the content authorship and management tool. With the exception of a relaxation tool that is provided only to the control group, a researcher can swap all content out with no programming knowledge. Visual elements on the home screen of the application draw content from an access and HTML database and would require knowledge of these resources and a few hours of work to change the workflow. Thus, an intervention with the exact same patient workflow could be created with no additional programming but an intervention with similar needs (eg, providing videos and PowerPoints to patients and collecting assessment data) could be constructed with a few hours of programming time.

Figure 2. Purple "editor" content authorship and management tool.
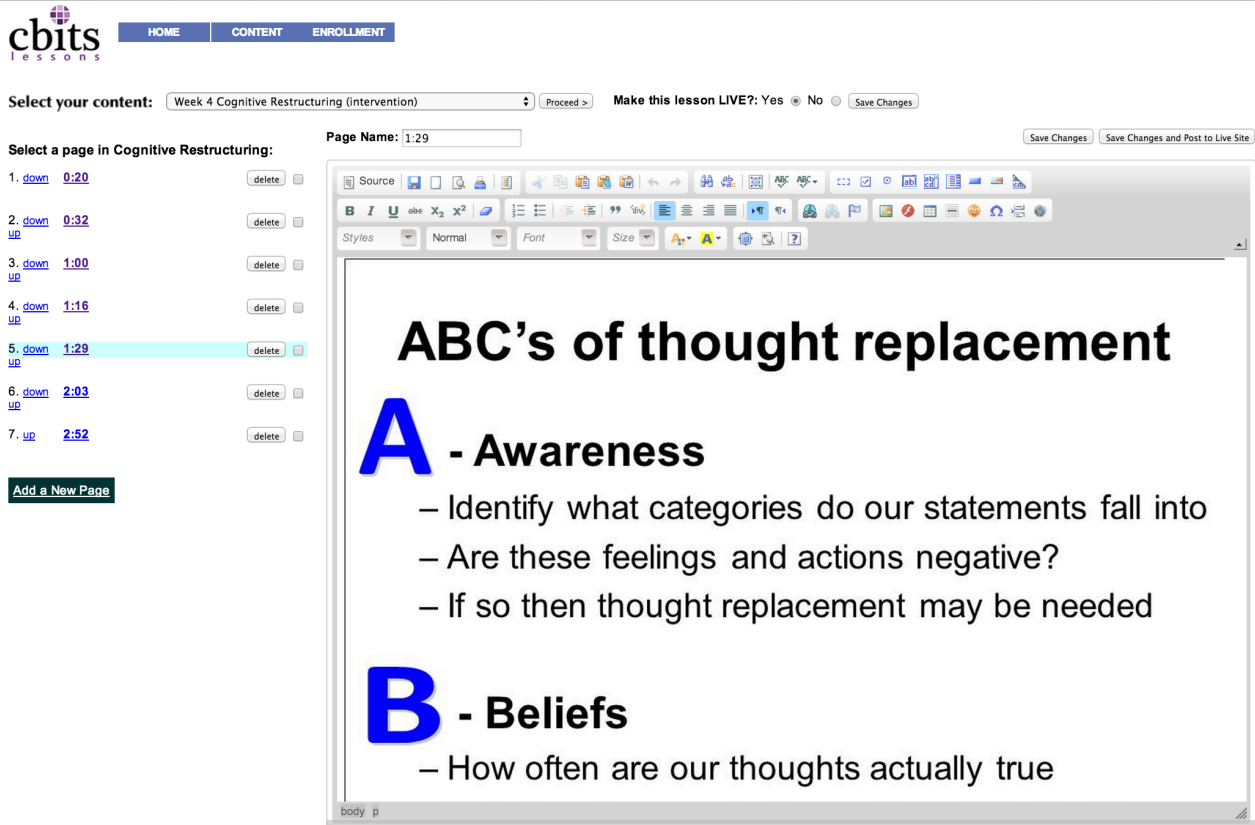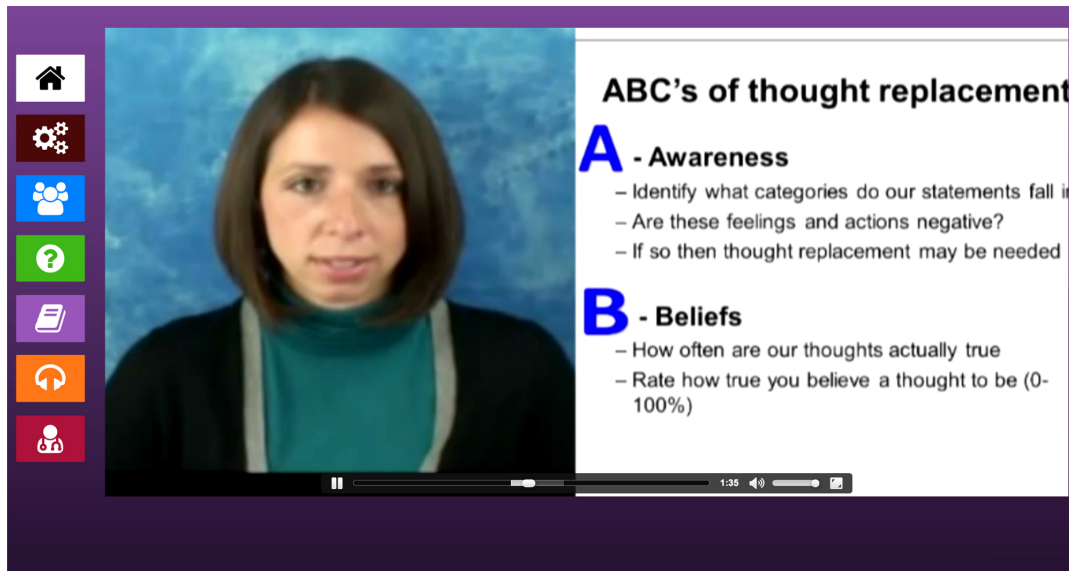

Figure 3. Patient-facing intervention constructed using content authorship and management tool.

The Future of Purple

Strengths and Limitations

We have presented the Purple framework and the conceptual issues related to its design and implementation as an example for the field of a modular and extensible system. The future of Purple will involve both the expansion of new features, greater translation of currently non-generalizable into generalizable features, and expansion of the research teams using and contributing to the Purple system. Although we have provided examples of BITs within Purple to show how developed features combine into a single intervention, we want to caution creators against conceptualizing and developing their interventions only at the feature-level.

One major challenge facing any development framework is to remain flexible and agile in face of the rapidly shifting technological environment. Currently, over 11,000 distinct devices exist that use the Android operating platform [26] and 23 that run iOS [27]. Furthermore, although these are currently the dominant platforms, emerging devices (eg, smartwatches,

Google Glass) might significantly shift the development environment. Software updates occur even more rapidly and change the requirements and capabilities of applications running on new versions of operating systems.

Fortunately, Purple is designed to be flexible in the face of these changes. However, its ability to adapt is tied to the different contexts that participating projects provide. While investigators in the community can benefit from the investments in software and knowledge provided from those who have used Purple before them, they also contribute to the system as their projects provide new challenges, problems, and opportunities. As solutions for these problems are developed within Purple, they in turn can be passed on to the community of researchers, either through the Purple system, or when components can exist independent of the system, such as Purple Robot [17], through opensourcing to the community. As it currently stands, Purple offers a collection of features and tools (user management, content authorship, data management, and phone sensor acquisition) that do not have to be recreated by outside technologists if they want to create projects with similar aims. The future of Purple rests on its ability to guide, both conceptual and practically, the development of the next generation of BITs.

Conclusions

The Purple system is an extensible, modular development architecture designed to support the creation of Web-based and mobile applications aimed at supporting behavior change. We have described the Purple system, focusing on the rationale guiding its development and the features available in its current state. We have demonstrated its use through a discussion of two ongoing projects investigating BITs for diverse treatment targets and populations. Last, we have presented guiding principles for the field to advance the way creators think about BITs and encourage further collaboration with the Purple framework.

While BITs hold great promise, actualizing this potential requires the capacity to transform behavior change principles into deployable interventions. Purple, now being used by more than 50 projects, facilitates this translation and does so in a way that allows for the repurposing of features and easier sharing of knowledge gained from evaluations. Purple is also a research platform that, through sharing of resources, increases capacity as well as the generalizability of knowledge.

Conflicts of Interest

None declared.

References

1. 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;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
2. Schueller SM, Munoz RF, Mohr DC. Realizing the potential of behavioral intervention technologies. Current Directions in Psychological Science 2013;22(6):478-483. [doi: 10.1177/0963721413495872]
3. Kazdin AE, Blase SL. Rebooting Psychotherapy Research and Practice to Reduce the Burden of Mental Illness. Perspectives on Psychological Science 2011;6(1):21-37. [doi: 10.1177/1745691610393527]
4. Vogt TM, Hollis JF, Lichtenstein E, Stevens VJ, Glasgow R, Whitlock E. The medical care system and prevention: the need for a new paradigm. HMO Pract 1998;12(1):5-13. [Medline: 10178378]
5. Stokols D, Hall KL, Taylor BK, Moser RP. The science of team science: Overview of the field and introduction to the supplement. Am J Prev Med 2008;35(2):S77-S89. [doi: 10.1016/j.amepre.2008.05.002]
6. Estrin D, Sim I. Health care delivery. Open mHealth architecture: an engine for health care innovation. Science 2010;330(6005):759-760. [doi: 10.1126/science.1196187] [Medline: 21051617]
7. Williams S, Yardley L, Wills GB. A qualitative case study of LifeGuide: users' experiences of software for developing Internet-based behaviour change interventions. Health Informatics J 2013 Mar;19(1):61-75. [doi: 10.1177/1460458212458915] [Medline: 23486826]
8. Fogg BJ. The behavior wizard. URL: http://www.behaviorwizard.org/wp/ [accessed 2014-06-17] [WebCite Cache ID 6QPRCkNAB]
9. Martin RWT. Agile software development: principles, patterns, and practices. Upper Saddle River, N.J.: Pearson Education; 2003.
10. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
11. Mohr DC, Duffecy J, Jin L, Ludman EJ, Lewis A, Begale M, et al. Multimodal e-mental health treatment for depression: a feasibility trial. J Med Internet Res 2010;12(5):e48 [FREE Full text] [doi: 10.2196/jmir.1370] [Medline: 21169164]

XSL•FO

RenderX

12. Duffecy J, Sanford S, Wagner L, Begale M, Nawacki E, Mohr DC. Project onward: an innovative e-health intervention for cancer survivors. Psychooncology 2013;22(4):947-951 [FREE Full text] [doi: 10.1002/pon.3075] [Medline: 22438297]

13. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]

14. Ben-Zeev D, Kaiser SM, Brenner CJ, Begale M, Duffecy J, Mohr DC. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. Psychiatr Rehabil J 2013;36(4):289-296. [doi: 10.1037/prj0000019] [Medline: 24015913]

15. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]

16. Olkin I. Statistical and theoretical considerations in meta-analysis. J Clin Epidemiol 1995;48(1):133-46; discussion 147. [Medline: 7853039]

17. Karr CJ. Purple robot. URL: http://tech.cbits.northwestern.edu/purple-robot/ [accessed 2014-03-04] [WebCite Cache ID 6NpXcDZqO]

18. Brown CH, Mohr DC, Gallo CG. A computational future for preventing HIV in minority communities: How advanced technology can improve implementation of effective programs. J Acquir Immune Defic Syndr 2013;63(Sup1):S72-S84. [doi: 10.1097/QAI.0b013e31829372bd]

19. 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]

20. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236. [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]

21. Ruotsalainen PS, Blobel BG, Seppälä AV, Sorvari HO, Nykänen PA. A conceptual framework and principles for trusted pervasive health. J Med Internet Res 2012;14(2):e52 [FREE Full text] [doi: 10.2196/jmir.1972] [Medline: 22481297]

22. Andersson G, Estling F, Jakobsson E, Cuijpers P, Carlbring P. Can the patient decide which modules to endorse? An open trial of tailored internet treatment of anxiety disorders. Cogn Behav Ther 2011;40(1):57-64. [doi: 10.1080/16506073.2010.529457] [Medline: 21337215]

23. Kanter J, Busch A, Rusch L. Behavioral activation: distinctive features. New York, NY: Routledge; 2009.

24. Penedo FJ, Antoni MH, Schneiderman N. Cognitive-behavioral stress management for prostate cancer recovery workbook. New York: Oxford University Press; 2008.

25. Assessment Center. URL: http://www.assessmentcenter.net/ [accessed 2014-03-04] [WebCite Cache ID 6NpYNMSlh]

26. OpenSignal. Android fragmentation visualized URL: http://opensignal.com/reports/fragmentation-2013/ [accessed 2014-02-28] [WebCite Cache ID 6NiDrCUx8]

27. Wikipedia. List of iOS devices URL: http://en.wikipedia.org/wiki/List_of_iOS_devices [accessed 2014-02-28] [WebCite Cache ID 6NiDwU0Tq]

Abbreviations

API: application programming interfaces
BIT: behavioral intervention technology
GPS: global positioning system
NIH: National Institutes of Health
PHI: personal health information
PROMIS: patient reported outcomes measurement information system

XSL•FO

RenderX

Original Paper

The Use of Social Media by State Tobacco Control Programs to Promote Smoking Cessation: A Cross-Sectional Study

Jennifer C Duke[1], PhD; Heather Hansen[1], MPP; Annice E Kim[1], PhD; Laurel Curry[1], MPH; Jane Allen[1], MA

RTI International, Public Health Policy Research, Research Triangle Park, NC, United States

Corresponding Author:
Jennifer C Duke, PhD
RTI International
Public Health Policy Research
3040 E. Cornwallis Road
PO Box 12194
Research Triangle Park, NC, 27709
United States
Phone: 1 919 485 2669
Fax: 1 919 541 6683
Email: jduke@rti.org

Abstract

Background: The promotion of evidence-based cessation services through social media sites may increase their utilization by smokers. Data on social media adoption and use within tobacco control programs (TCPs) have not been reported.

Objective: This study examines TCP use of and activity levels on social media, the reach of TCP sites, and the level of engagement with the content on sites.

Methods: A cross-sectional descriptive study of state TCP social media sites and their content was conducted.

Results: In 2013, 60% (30/50) of TCPs were using social media. Approximately one-quarter (26%, 13/50) of all TCPs used 3 or more social media sites, 24% (12/50) used 2, and 10% (5/50) used 1 site. Overall, 60% (30/50) had a Facebook page, 36% (18/50) had a Twitter page, and 40% (20/50) had a YouTube channel. The reach of social media was different across each site and varied widely by state. Among TCPs with a Facebook page, 73% (22/30) had less than 100 likes per 100,000 adults in the state, and 13% (4/30) had more than 400 likes per 100,000 adults. Among TCPs with a Twitter page, 61% (11/18) had less than 10 followers per 100,000 adults, and just 1 state had more than 100 followers per 100,000 adults. Seven states (23%, 7/30) updated their social media sites daily. The most frequent social media activities focused on the dissemination of information rather than interaction with site users. Social media resources from a national cessation media campaign were promoted infrequently.

Conclusions: The current reach of state TCP social media sites is low and most TCPs are not promoting existing cessation services or capitalizing on social media's interactive potential. TCPs should create an online environment that increases participation and 2-way communication with smokers to promote free cessation services.

Introduction

Approximately 18% of US adults are cigarette smokers [1]. Evidence from the National Health Interview Survey indicates that, in 2010, 52% of adult smokers in the United States tried to quit in the past year, but only 6% succeeded [2]. The low success rate for smokers' quit attempts may be due, in part, to the low proportion (31%) of smokers who used evidence-based interventions as part of their quit attempt [2]. All US state tobacco control programs (TCPs) currently offer free evidence-based smoking cessation services through telephone quitlines, and half also offer Web-based interventions. However, the paid promotion of cessation services is limited because of low levels of state TCP funding [3]. Recent national promotion of cessation services has sporadically increased cessation service utilization for brief time periods since 2012 [4]. Given that the decline in the prevalence of smoking has slowed in recent years [5], sustained and innovative approaches are needed to increase

XSL•FO
RenderX

the promotion, utilization, and reach of these interventions to maximize their effectiveness [6-8].

The promotion of evidence-based cessation services through social media sites may be a low-cost strategy to increase their utilization by smokers. Almost 3 in 4 US adults aged 18 years or older use at least 1 social media site [9]. Recent data indicate that 71% of online adults use Facebook and 63% of Facebook users visit the site daily [9]. In 2013, 18% of online adults used Twitter and 8% used Twitter daily [9]. Facebook is widely used by online adults across all age groups (84% of adults aged 18 to 29, 79% of adults aged 30 to 49, 60% of adults aged 50 to 64, and 45% of adults aged 65 or older), whereas online adults using Twitter tend to skew to a younger demographic (31% of adults aged 18 to 29, 19% of adults aged 30 to 49, 9% of adults aged 50 to 64, and 5% of adults aged 65 or older) [9]. Photo-sharing sites are also increasing in popularity; for example, 21% of US adults currently use Pinterest and 17% use Instagram [9]. In addition to linking smokers to existing cessation services, social media sites provide a venue for relevant, credible tobacco cessation messages to reach smokers and their social circles. Studies show that people are using social media to seek health information, share their health experiences, and provide and receive psychological support [10-12]. Social media sites also expand social networks, which may influence behavioral and emotional change [13-15]. Media messages promoting free smoking cessation interventions may be shared among smokers, their families, and their friends, motivating quit attempts.

In keeping with the growth of social media activities in the United States, many states currently use social media sites to disseminate health information [16,17]. The most recent descriptive study found that 82% (41/50) of all state public health departments use Twitter and 56% (28/50) use Facebook [18]. A second study of state health departments found that the reach of health department sites relative to the states' populations was low and, in most cases, the sites failed to take advantage of the fundamental advantage of this medium: engaging users in interactive communication [19]. Harris and colleagues [20] found that state health departments with social media sites were located in more populated states and had higher per capita health department expenditures compared to those without social media sites.

Recently, state TCPs gained new access to free cessation-related content and promotions of cessation services specifically designed to be shared through social media sites. In 2013, the Centers for Disease Control and Prevention (CDC) aired the second wave of the first federally funded tobacco education campaign, Tips From Former Smokers (*Tips*), which encourages quitting through advertisements that show personal testimonials from ex-smokers battling life-altering smoking-related illnesses. The campaign included a large social media presence; CDC shared *Tips*-related materials and social media links publicly on its website and encouraged the public and practitioners to share its social media content with their networks. CDC's Facebook and Twitter accounts posted frequently during the campaign, sharing updates and campaign-related information, videos, and images.

This study provides a descriptive overview of state TCP social media sites and the extent to which they are used to disseminate cessation messages and promote quitting services. The purpose of this study is to document the use of state TCP social media sites during 2013, including Facebook, Twitter, YouTube, and others. The extent to which TCPs actively use their social media sites, the reach of TCP sites, and audience interaction and engagement with TCP social media sites are examined. This study also explores the content of messages posted on social media by TCPs, particularly state TCPs' use of existing social media material from the national *Tips* campaign during the months it aired in 2013.

Methods

Sample and Data Collection

To identify state TCPs engaged with social media sites, we conducted an online Google search of each state's name and the keywords "tobacco program" or "quitline." We then reviewed the websites for hyperlinks or information on related social media sites, including Facebook, Twitter, YouTube, Pinterest, Instagram, Google+, blogs, and virtual communities. Virtual communities are social networks of individuals who interact online; TCP-sponsored communities that encourage smokers to support one another during quit attempts are included in this study. If no links existed, the quitline service name (eg, Quit Now Kentucky) was used in a secondary Google search to identify social media sites. We also included any tobacco-related social media sites sponsored by a state health department but not obviously affiliated with the TCP or quitline name (eg, Make Smoking History, KanQuit!). The search identified 30 states with Facebook sites (ie, pages), 18 states with Twitter sites (ie, accounts identified by name or "handle"), and 20 YouTube sites (ie, channels) as of December 31, 2013 (Multimedia Appendix 1). Two states had more than 1 social media site on the same platform (eg, 2 Facebook pages), in which case the site focused on tobacco cessation was selected for analysis. Publicly available activity data and metrics on all social media sites were collected using Snagit screen capture software (TechSmith Corporation, Okemose, MI, USA). In addition, the social media monitoring tool radian6 (Salesforce Marketing Cloud, New York, NY, USA) was used to capture posted content by users on Twitter.

Content Coding

The content of state TCP posts to social media sites was collected and analyzed for the subset of months in 2013 during which the national *Tips* smoking cessation campaign aired (March to June 2013). Using an inductive approach, coders observed a subset of messages to develop the coding scheme [21]. Sites with no TCP activity during this time frame were excluded from the content analysis, for a total of 25 Facebook pages, 16 Twitter accounts, and 16 YouTube channels available for coding. Three researchers independently coded 1123 Facebook posts, 1776 Twitter tweets, and 591 YouTube videos (eg, number of likes, shares, comments, message content). Content coding of the most popular social media sites confirmed their primary communication focus: 70.5% (1252/1776) of the content on Twitter and 67.3% (756/1123) of the content on

Facebook was about tobacco cessation (and other topics included secondhand smoke and tobacco control policies). Interrater reliability among coders on 5% of data was 95.7% or higher for all coding categories (kappa=.72).

Measures

Overview

Measurement techniques used to guide social media analysis and management that have been pioneered in the for-profit sector [22] may be applied to the study of state TCP social media sites. Key aspects of social media evaluation include exposure (reach or the number of people reached with a message), engagement (number of people who take action in response to a message), and influence (whether engagement is positive, neutral, or negative in sentiment) [23,24]. This study assessed 2 of these 3 evaluation components: exposure to and engagement with state TCP social media sites. Study measures described herein include the presence of social media sites, TCP site activity, audience reach, audience engagement, and social media site promotional activities.

Tobacco Control Program Presence on Social Media

We noted whether TCPs had a presence on social media sites (eg, Facebook, Twitter, YouTube) as of December 31, 2013, and the date of site establishment.

Tobacco Control Program Social Media Activity

Site activity was defined as the volume of content posted by the TCP on each social media site (ie, the number of posts, tweets, videos) from January to December 2013.

Audience Reach

Audience reach on TCP social media sites was measured by Facebook page likes, Twitter followers, and YouTube views (which may represent more than 1 person, but no individual metric is available) as of December 31, 2013. For example, on Facebook, users may "like" a TCP page, enabling posts by the TCP to be sent to users' newsfeeds where they are the most likely to be seen. Reach metrics are presented as a total number and also as a proportion of the state adult population. State adult population was calculated using 2010 census adult (aged 18 or older) population data and is reported as reach per 100,000 adults.

Audience Engagement

Audience interaction and exchange of messages were assessed across social media sites from March to June 2013. As indicators of action on the part of the audience, they serve as a proxy for engagement with the audience. User activity generates content that may be viewed by other users, thus expanding the reach of the site's messages. To measure audience engagement, data on the amount and relative proportion of the following variables were collected: (1) individual post likes, comments, and shares on Facebook; (2) retweets, defined as the number of times a TCP post (ie, tweet) was shared by another Twitter user; and (3) video likes, shares, and comments on YouTube [19,23].

Promotion of State Cessation Services and the Tips Campaign

Coders documented the presence of a link/URL to the states' cessation resources (ie, the state quitline phone number or website) in each post on Facebook or Twitter from March to June 2013. All posts during this time frame were also coded for the presence of CDC *Tips* campaign content. This included links to *Tips* campaign materials (eg, television advertisements), online campaign stories of the smokers in the *Tips* ads, and shares of CDC's posts or tweets about *Tips*. YouTube videos were not coded because they were primarily promoted or linked to Facebook and/or Twitter, where the information was captured and coded.

Photo and Video Content

Because research has shown that content with images and video attracts more views [25,26], the presence of either photos or videos for each Facebook or Twitter post from March to June 2013 was coded.

Analysis

We report the results of the collected metrics and content analysis using simple descriptive statistics, including measures of central tendency, frequencies, and counts in summary charts. Because of the limited presence on many social media sites examined in the study, most measures are reported for the 3 most popular sites: Facebook, Twitter, and YouTube.

Results

Presence of Social Media Sites

Of all TCPs, 60% (30/50) used social media in 2013. Nearly one-quarter (26%, 13/50) of all TCPs used 3 or more social media sites, 24% (12/50) used 2 sites, and 10% (5/50) used 1 site. The use of Facebook was most common, with 60% (30/50) of state TCPs using the site. Twitter was used by 36% (18/50) of state TCPs and 40% (20/50) had a YouTube channel. In addition to Facebook, Twitter, and YouTube, few state TCPs offered other social media sites. Three states used Pinterest (Louisiana, Mississippi, and Oregon), 2 states had their own blog (California and New Hampshire), and 1 state offered Google+ (New York). In addition, 3 states provided a virtual community site for smoking cessation (Arizona, Connecticut, and New York). No states used Instagram in 2013.

Tobacco Control Program Site Activity

Activity was examined on the 3 most popular social media sites: Facebook, Twitter, and YouTube. Figures 1 and 2 show the average monthly frequency with which TCPs posted content to their Facebook site and disseminated content via Twitter from January to December 2013.

Of the 30 states with Facebook pages, most (22/30, 73%) posted content (ie, messages, photos, videos) an average of 6 to 9 times per month. Two states, Florida and Illinois, posted content 10 times or more per month on average (Florida: mean 61, SD 12; Illinois: mean 14, SD 9). The median number of yearly Facebook posts was 89 (range 0-727). Among all TCP posts

coded for content from March to June 2013, 39.70% (705/1776) included photos and 7.66% (136/1776) included videos.

Fewer state TCPs used Twitter than Facebook. However, those that used Twitter delivered more content through Twitter than Facebook. On average, Florida posted content most often at 220 times per month and 6 other states posted at least once per day. The median number of yearly TCP tweets was 118 (range

15-2618). Among all TCP tweets coded, 10.69% (120/1123) included links to photos and 6.06% (68/1123) included links to videos.

In 2013, the median number of videos posts to TCP YouTube sites was 39 (range 3-276). Videos consisted primarily of uploaded antitobacco television advertisements.

Figure 1. Average number of monthly Facebook posts by state tobacco control programs (n=30), January to December 2013. Gray lines denote standard deviation of monthly posts.
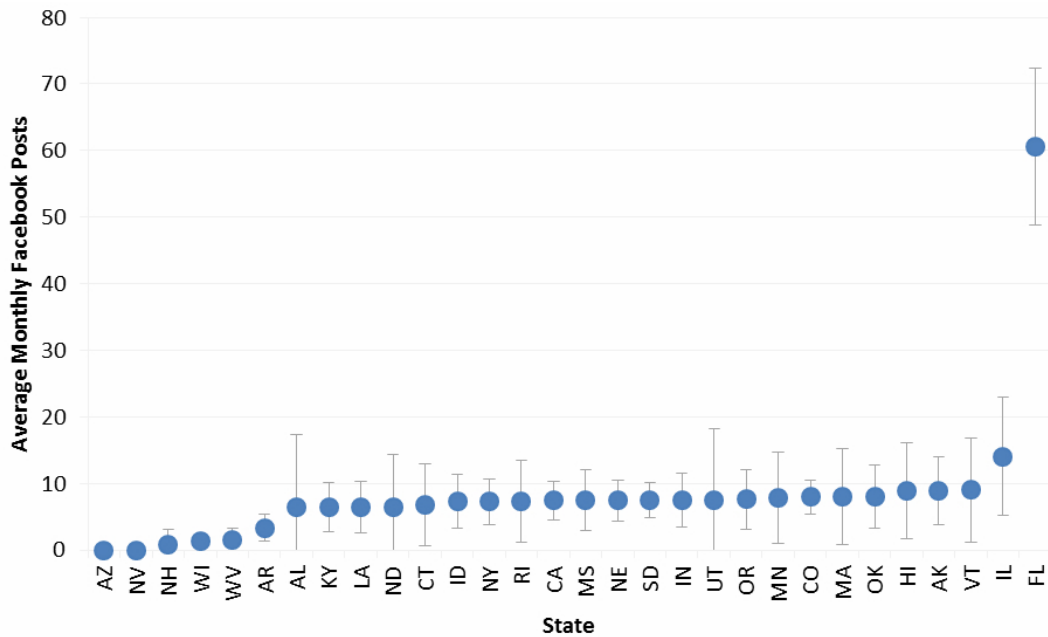
Figure 2. Average number of monthly tweets by state tobacco control programs (n=17), January to December 2013. Gray lines denote standard deviation of monthly tweets. Data for AZ are not included in the graph because there was no posted content to its Twitter site in 2013.
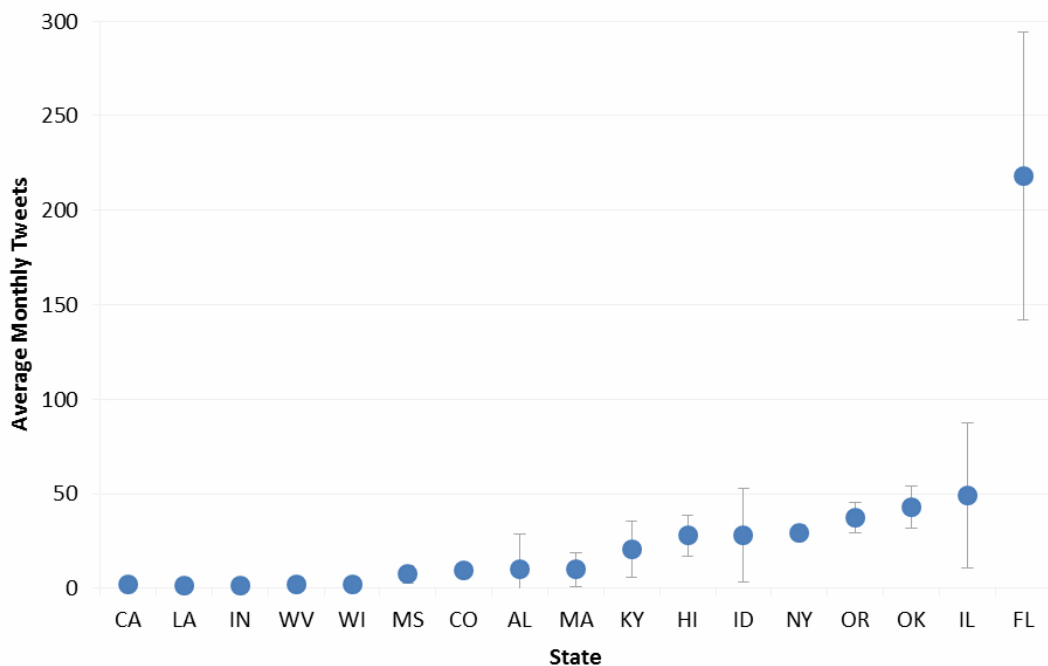
Audience Reach

The reach of social media sites varied widely by platform and by state in 2013 (Table 1). Overall, the level of exposure to

social media sites was low relative to states' populations. The median number of people who liked a TCP Facebook site was 893; however, the range was 72 to 226,722, suggesting dramatic variation in reach. Among TCPs with a Facebook page, 73%

(22/30) had less than 100 likes per 100,000 adults in the state, and 13% (4/30) had more than 400 likes per 100,000 adults. Overall, the median number of Twitter followers was lower than Facebook likes at 122 (range 13-2690). Among TCPs with a Twitter page, 61% (11/18) had less than 10 followers per 100,000 adults and only 1 state (Hawaii) had more than 100 followers per 100,000 adults. The median number of views of TCP videos on YouTube was 5216 (range 34-1,308,248); however, there was enormous variation across TCPs. For example, the yearly video views of Florida's YouTube channel alone accounted for 60% of all video views for the 20 state TCP YouTube sites.

Audience Engagement

There was considerable variation in the level of audience engagement on TCP social media sites from March to June 2013 (Table 2). The 4 TCPs that generated the most audience engagement through Facebook were California, with a median of 114 (range 5-1370) engagement activities per post (ie, through post likes, comments, or shares), Florida (median 116, range 5-1277), South Dakota (median 53, range 10-461), and Alaska (median 33, range 0-261). In addition to these 4 states, posts in a few other states had high levels of audience interaction. For example, 81% (25/31) of Minnesota's TCP Facebook posts were commented on. Similarly, Kentucky, New York, Oklahoma, and Utah had at least 1 like on 90% or more of their posts. Across all states, the median of 27% (range 0-95) of Facebook posts had at least 1 comment.

Compared with Facebook users, Twitter users were less likely to share content with their social circle. On average, nearly 1 in 10 TCP tweets (9.68%, 537/5547) were retweeted during the study period. In 2013, Florida's Twitter content was shared with others most frequently (40.56%, 1062 of 2618 TCP tweets retweeted), followed by Oklahoma (23.1%, 119/515). Another 7 states had between 20 and 65 TCP tweets shared with others in 2013 (Hawaii, Idaho, Illinois, Kentucky, Massachusetts, New York, and Oregon), and the remaining 8 states on Twitter had fewer than 10 TCP retweets (Figure 2). Across TCP YouTube sites, less than one-third of all posted videos (30.5%, 180/591) received at least 1 like. Also, 1 in 6 videos (15.6%, 92/591) received 1 comment or more from viewers.

Promotion of State Cessation Services and the Tips Campaign

During the airing of the *Tips* campaign from March to June 2013, 64% (16/25) of TCP Facebook sites had 5 or fewer posts that promoted cessation services. Across all sites, approximately 15% (14.60%, 164/1123) of all TCP Facebook posts included materials such as television advertisements or links to CDC's *Tips* campaign. In total, 64.11% (720/1123) of all TCP Facebook posts coded included no promotion of either telephone quitlines or Web-based cessation services available to smokers.

Of the TCP Twitter sites in use over the 4-month period of *Tips*, 56% (9/16) had 6 or fewer tweets that promoted cessation services. In total, 17.85% (317/1776) of all Twitter content included materials such as television advertisements or links to CDC's *Tips* campaign while it was on air. On average across states, 75.00% (1332/1776) of all TCP Twitter content included no promotion of either telephone quitlines or Web-based cessation services available to smokers.

Table 1. Date of site establishment and reach metrics for TCP social media sites (N=30).

State	Facebook			Twitter			YouTube		
	Page start date	Number of page likes	Likes per 100,000 adults	Date of first tweet	Followers	Followers per 100,000 adults	Channel start date	Total views	Total videos
AL	8/20/09	411	11.2	4/23/10	77	2.1	—	—	—
AK	6/14/10	2257	422.4	—	—	—	—	—	—
AZ	6/4/10[a,b]	512	10.5	6/27/12	13	0.3	4/26/12	4170	58
AR	2/16/10	848	38.1	—	—	—	1/24/08	43,287	277
CA	12/1/09	59,309	208.7	3/8/13	18	0.1	12/9/09	470,835	28
CO	6/21/11	787	20.2	6/11/12	96	2.5	4/11/13	616	14
CT	7/26/11	2328	83.8	—	—	—	7/26/11	34	6
FL	1/28/08	226,722	1505.1	3/29/10	2690	17.9	1/28/08	1,308,248	60
HI	12/30/10	4686	437.9	7/28/09	1688	157.7	7/6/09	6262	36
ID	5/14/09	1480	127.9	9/1/09	519	44.9	—	—	—
IL	10/29/13	1313	13.4	10/29/13	59	0.6	—	—	—
IN	12/9/09	1672	34.0	7/28/11	170	3.5	12/9/09	1215	4
KY	9/16/11	526	15.7	9/20/11	100	3.0	12/21/11	1199	3
LA	1/14/11	547	15.8	2/14/11	377	10.9	1/13/12	65	1
MA	4/8/09	937	18.1	1/26/11	620	12.0	3/5/09	14,525	62
MN	8/5/09	2421	59.5	—	—	—	11/1/09	44,993	41
MS	7/7/11	130	5.8	7/8/11	80	3.6	5/12/11	686	23
NE	5/16/11	1086	78.5	—	—	—	—	—	—
NV	10/26/10	83	4.0	—	—	—	—	—	—
NH	5/23/11	408	39.3	—	—	—	7/26/10	1243	13
NY	7/15/09	805	5.3	11/25/09	143	0.9	6/18/10	20,863	17
ND	8/27/13	72	13.5	—	—	—	1/23/12	2118	3
OK	11/22/11	2658	93.1	11/12/12	329	11.5	11/21/13	146	1
OR	12/31/09	8416	279.8	1/4/10	1369	45.5	9/21/11	80,578	8
RI	12/12/12[a]	88	10.6	—	—	—	—	—	—
SD	1/10/11	9344	1504.8	—	—	—	—	—	—
UT	3/23/11	740	38.2	—	—	—	5/31/11	34,557	81
VT	11/2/09	1417	283.2	—	—	—	12/28/12	140,324	19
WV	3/16/10	611	41.5	3/16/10	37	2.5	—	—	—
WI	4/26/10	159	3.6	8/13/09	70	1.6	—	—	—
Median	6/14/10	893	38	9/9/10	122	3	5/21/11	5216	18
Min	1/28/08	72	4	7/28/09	13	0	1/24/08	34	1
Max	10/29/13	226,722	1505	10/29/13	2690	158	11/21/13	1,308,248	277

[a] No start date recorded; date of first post.

[b] Last post August 23, 2012.

XSL•FO
RenderX

Table 2. Audience engagement with TCP Facebook sites and content of posts, March to June 2013.

State[a]	TCP posts, n	Audience interaction and engagement				Promotional posts, n (%)		Post features, n (%)	
		Audience activity[b] per post, median (range)	Posts, n (%)			Tips campaign	State cessation services[c]	Posts with video	Posts with photo
			With likes	With shares	With comments				
AL	46	1 (0-9)	15 (33)	11 (24)	5 (11)	8 (17)	10 (22)	4 (9)	24 (52)
AK	58	33 (0-261)	56 (97)	36 (62)	40 (69)	2 (3)	32 (55)	0 (0)	52 (90)
AR	12	2 (0-10)	7 (58)	2 (17)	4 (33)	2 (17)	2 (17)	0 (0)	1 (8)
CA	48	114 (5-1370)	48 (100)	45 (94)	45 (94)	6 (13)	10 (21)	3 (6)	28 (58)
CO	44	3 (0-87)	31 (71)	12 (27)	6 (14)	11 (25)	13 (30)	5 (11)	24 (55)
CT	48	3 (0-40)	37 (77)	11 (23)	16 (33)	0 (0)	1 (2)	0 (0)	2 (4)
FL	231	116 (5-1277)	231 (100)	229 (99)	219 (95)	8 (4)	94 (41)	13 (6)	134 (58)
HI	28	2 (0-7)	20 (71)	1 (4)	5 (18)	1 (4)	6 (21)	0 (0)	14 (50)
ID	44	1 (0-14)	27 (61)	6 (14)	3 (7)	0 (0)	3 (7)	2 (5)	8 (18)
IN	60	4 (0-23)	48 (80)	48 (80)	16 (27)	16 (27)	38 (63)	11 (18)	39 (65)
KY	38	7 (0-27)	35 (92)	11 (29)	8 (21)	1 (3)	0 (0)	3 (8)	16 (42)
LA	17	2 (0-15)	13 (77)	7 (41)	3 (18)	9 (53)	5 (29)	2 (12)	7 (41)
MA	35	4 (0-11)	27 (77)	15 (43)	10 (29)	2 (6)	2 (6)	1 (3)	25 (71)
MN	31	10 (0-107)	30 (97)	16 (52)	25 (81)	0 (0)	17 (53)	6 (19)	3 (10)
MS	41	1 (0-5)	15 (37)	10 (24)	2 (5)	30 (73)	2 (5)	5 (12)	28 (68)
NE	42	4 (0-102)	35 (83)	8 (19)	16 (38)	2 (5)	17 (41)	1 (2)	23 (55)
NY	38	5 (0-11)	36 (95)	14 (37)	9 (24)	0 (0)	19 (50)	1 (3)	3 (8)
OK	47	14 (1-120)	47 (100)	21 (45)	29 (62)	5 (11)	31 (66)	6 (13)	18 (38)
OR	50	6 (0-41)	40 (80)	17 (34)	24 (48)	8 (16)	20 (40)	2 (4)	17 (34)
RI	50	2 (0-9)	30 (60)	14 (28)	3 (6)	6 (12)	13 (26)	4 (8)	8 (16)
SD	31	53 (10-461)	31 (100)	28 (90)	25 (81)	5 (16)	18 (58)	2 (7)	25 (81)
UT	54	12 (4-215)	54 (100)	51 (94)	15 (28)	0 (0)	43 (80)	44 (82)	1 (2)
VT	16	1 (0-14)	6 (38)	7 (44)	2 (13)	3 (19)	3 (19)	3 (19)	4 (25)
WV	7	2 (0-17)	5 (71)	3 (43)	0 (0)	1 (14)	2 (29)	1 (14)	3 (43)
WI	7	1 (0-3)	4 (57)	2 (29)	0 (0)	2 (29)	2 (29)	3 (43)	0 (0)
Median	42	4 (0-1370)	31 (77)	12 (37)	9 (27)	2 (12)	10 (29)	3 (8)	16 (42)

[a] Five states (Arizona, Illinois, Nevada, New Hampshire, and North Dakota) with Facebook sites did not post content during the time period.

[b] Number of individual post likes, comments, or shares.

[c] Post contains the state or national telephone quitline number and/or Web-based cessation services.

Discussion

Although 30 state TCPs adopted at least 1 form of social media, only 7 disseminated social media messages on a daily basis in 2013. Most content on TCP social media sites did not promote cessation services. Moreover, few state TCPs take advantage of the social media features that generate the most audience engagement—sharing photos and videos—or the opportunity to link to related content, such as the CDC's *Tips* campaign.

Perhaps as a result of these social media practices, most TCPs do not generate a substantial amount of audience engagement, limiting the value of potentially dynamic, interactive communication resources in tobacco control.

Our results corroborate other research indicating that the application of social media to public health research and practice has not taken advantage of its potential for multidirectional communication [27,28]. Study findings suggest that most state TCPs could benefit from strategic communication plans for

social media to increase reach and encourage audience interaction and engagement [16]. The Florida TCP, the most prolific user of social media in 2013, presents an example of how effective use of social media can generate audience engagement. During the study period, Florida's TCP promoted an average of 11 daily posts to its sites, and followers of the Florida TCP retweeted 41% of all content posted on Twitter. On average, there were more than 185 interactions for each posting to Florida's Facebook site (ie, people who liked, commented on, or shared content), expanding the reach of each individual message. The most successful TCP social media sites communicated information in a way that reflected the audience preferences (eg, with videos and photos), thereby encouraging audience response and discussion. Without engaging in interactive communication with their audience, state TCPs are unable to characterize their audience or receive feedback for site improvements [28].

Taken together, these findings indicate that changes in the social media practices of state TCPs can enhance their ability to connect with constituents—including smokers who are interested in quitting—in ways that may promote cessation. State TCPs have a long history of providing evidence-based cessation services. Although improvements to the content and interactive nature of TCP social media sites are contingent upon their ability to dedicate human resources to site management, states have access to well-designed tobacco cessation online materials from a variety of government and nonprofit agencies (eg, CDC's Office on Smoking and Health, Public Health Service, National Cancer Institute). To encourage participation, TCPs should focus on the dissemination of evidence-based tobacco cessation information in an engaging format promote appropriate behavioral interventions, and create an online environment that increases participation and 2-way communication with users.

Active TCP social media sites could influence how smokers discover, access, and understand cessation- and tobacco-related information, including connecting the public to telephone quitlines and Web-based cessation services. Traditional media campaigns are often most effective when accompanied by access to available services [29]. Social media messages can reach consumers with both information and links to services, such as telephone quitlines. In addition, social media sites present a unique opportunity to address state TCP goals for increasing cessation because they have a very low barrier for affiliation and present smokers with the option to join large ad hoc networks within their state [30]. TCP-driven content and interactions with users should be aimed at increasing smokers' use of evidence-based treatment options and enhancing self-efficacy during the quitting process.

As private and government organizations move toward online technologies to promote and provide services, there is a greater need for public health practitioners to be able to create evidence-based websites and promote strategies that will maximize exposure to evidence-based cessation services [31]. State TCPs face a wide range of barriers, including lack of dedicated staff within state health departments to encourage improvements to social media, lack of knowledge regarding best practices in social media, lengthy approval processes or policy restrictions that hinder improvements, and financial limitations for cessation-related promotional activities and services [3,29]. Large-scale and rapid improvements to state TCP social media sites may be bolstered by active collaboration among state and federal agencies, both of whom seek to improve population health outcomes by reducing smoking prevalence.

Although the emergence of social media sites presents a unique opportunity for state TCPs to communicate and engage with smokers, literature is lacking to evaluate the effectiveness of these channels in promoting cessation services. Some evidence suggests that online interaction and connectedness can increase cessation self-efficacy [32]. Wide-scale adoption of improvements to the content and interactive nature of TCP social media sites may encourage supportive communities of online smokers and yield increases in the use of evidence-based interventions. Social network analysis may help us better understand how federal and state TCPs are connected and how these networks can be leveraged to help inform more strategic communications plan. More experimental and applied research across multidisciplinary teams (eg, social media researchers, federal and state tobacco control cessation managers, and marketers) is necessary to understand the conditions under which TCP social media sites successfully engage and motivate smokers to quit.

This study has several limitations. Data could not be captured for messages that were deleted, noted as private by a TCP, or shared offline. Also, some TCPs may not have promoted state cessation services heavily through social media during the airing of the national media campaign, but may have done so during other time periods. Further, this study provides a quantitative summary of audience interaction and did not examine the type of site users or systematically examine users' posted content. Although state TCPs primarily target smokers in their campaigns, audiences following state TCPs may also include other health agencies [18]. Similarly, some TCP sites may be linked to cessation programs not intended for smokers (eg, California's Twitter). Future research on the type of site users and posted content should yield a better understanding of the level of engagement by smokers as well as public health organizations and other audiences [33]. Despite its limitations, this study provides an important first look at state TCP social media sites and audience engagement.

Conflicts of Interest

None declared.

Multimedia Appendix 1

State TCP Facebook, Twitter, and YouTube Links, December 2013.

[XLSX File (Microsoft Excel File), 23KB - jmir_v16i7e169_app1.xlsx]

References

1. Schiller J, Ward B, Freeman G. Centers for Disease Control and Prevention. 2013 Dec 17. Early release of selected estimates based on data from the January–June 2013 National Health Interview Survey URL: http://www.cdc.gov/nchs/nhis/released201312.htm [accessed 2014-07-01] [WebCite Cache ID 6Qk19Pzj7]
2. Centers for Disease ControlPrevention (CDC). Quitting smoking among adults--United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
3. Centers for Disease Control and Prevention. Best practices for comprehensive tobacco control programs. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014. URL: http://www.cdc.gov/tobacco/stateandcommunity/best_practices/pdfs/2014/comprehensive.pdf?utm_source=rss&utm_medium=rss&utm_campaign=best-practices-for-comprehensive-tobacco-control-programs-2014-pdf [WebCite Cache ID 6Qk1NYd3f]
4. Centers for Disease ControlPrevention (CDC). Increases in quitline calls and smoking cessation website visitors during a national tobacco education campaign--March 19-June 10, 2012. MMWR Morb Mortal Wkly Rep 2012 Aug 31;61(34):667-670 [FREE Full text] [Medline: 22932300]
5. US Department of Health and Human Services. The health consequences of smoking-50 years of progress: A report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2012. URL: http://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf [accessed 2014-07-01] [WebCite Cache ID 6Qk1eTMva]
6. Backinger CL, Thornton-Bullock A, Miner C, Orleans CT, Siener K, DiClemente CC, et al. Building consumer demand for tobacco-cessation products and services: The national tobacco cessation collaborative's consumer demand roundtable. Am J Prev Med 2010 Mar;38(3 Suppl):S307-S311. [doi: 10.1016/j.amepre.2009.12.002] [Medline: 20176301]
7. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
8. Cobb NK, Graham AL. Health behavior interventions in the age of facebook. Am J Prev Med 2012 Nov;43(5):571-572. [doi: 10.1016/j.amepre.2012.08.001] [Medline: 23079184]
9. Duggan M, Smith A. Social media update 2013. Washington, DC: Pew Internet & American Life Project; 2013 Dec 30. URL: http://www.pewinternet.org/2013/12/30/social-media-update-2013/ [accessed 2014-06-19] [WebCite Cache ID 6QS1KStGq]
10. De la Torre-Díez I, Díaz-Pernas FJ, Antón-Rodríguez M. A content analysis of chronic diseases social groups on Facebook and Twitter. Telemed J E Health 2012;18(6):404-408. [doi: 10.1089/tmj.2011.0227] [Medline: 22650380]
11. Keim-Malpass J, Steeves RH. Talking with death at a diner: young women's online narratives of cancer. Oncol Nurs Forum 2012 Jul;39(4):373-8, 406. [doi: 10.1188/12.ONF.373-378] [Medline: 22750895]
12. Sugawara Y, Narimatsu H, Hozawa A, Shao L, Otani K, Fukao A. Cancer patients on Twitter: a novel patient community on social media. BMC Res Notes 2012;5:699 [FREE Full text] [doi: 10.1186/1756-0500-5-699] [Medline: 23270426]
13. Christakis NA. Health care in a web. BMJ 2008 Jun 28;336(7659):1468 [FREE Full text] [doi: 10.1136/bmj.a452] [Medline: 18583678]
14. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. Am J Public Health 2010 Jul;100(7):1282-1289 [FREE Full text] [doi: 10.2105/AJPH.2009.165449] [Medline: 20466971]
15. Rosenquist JN, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. Ann Intern Med 2010 Apr 6;152(7):426-33, W141 [FREE Full text] [doi: 10.7326/0003-4819-152-7-201004060-00007] [Medline: 20368648]
16. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. BMC Public Health 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]
17. Emery S, Aly EH, Vera L, Alexander RL. Tobacco control in a changing media landscape: how tobacco control programs use the internet. Am J Prev Med 2014 Mar;46(3):293-296. [doi: 10.1016/j.amepre.2013.11.005] [Medline: 24512869]
18. Harris JK, Choucair B, Maier RC, Jolani N, Bernhardt JM. Are public health organizations tweeting to the choir? Understanding local health department Twitter followership. J Med Internet Res 2014;16(2):e31 [FREE Full text] [doi: 10.2196/jmir.2972] [Medline: 24571914]

XSL•FO
RenderX

19. Neiger BL, Thackeray R, Burton SH, Giraud-Carrier CG, Fagen MC. Evaluating social media's capacity to develop engaged audiences in health promotion settings: use of Twitter metrics as a case study. Health Promot Pract 2013 Mar;14(2):157-162. [doi: 10.1177/1524839912469378] [Medline: 23271716]

20. Harris JK, Snider D, Mueller N. Social media adoption in health departments nationwide: the state of the states. Frontiers in PHSSR 2013;2(1):5-5. [doi: 10.2105/AJPH.2012.301166]

21. Neuendorf KA. The Content Analysis Guidebook. Thousand Oaks, CA: Sage Publications; 2002.

22. Peters K, Chen Y, Kaplan AM, Ognibeni B, Pauwels K. Social media metrics — A framework and guidelines for managing social media. Journal of Interactive Marketing 2013 Nov;27(4):281-298. [doi: 10.1016/j.intmar.2013.09.007]

23. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]

24. The Conclave. Complete social media measurement standards. 2013. URL: http://www.smmstandards.com/wp-content/uploads/2013/06/Complete-standards-document4.pdf [accessed 2014-06-19] [WebCite Cache ID 6QS2PrKOq]

25. Cooper B. Fast Company. 2013. 7 powerful Facebook statistics you should know about Fast Company URL: http://www.fastcompany.com/3022301/work-smart/7-powerful-facebook-statistics-you-should-know-about [accessed 2014-06-19] [WebCite Cache ID 6QS2R7he7]

26. Neher K. Smart Business. 2013. Images versus text - which is stronger in the world of social media marketing? URL: http://www.sbnonline.com/component/k2/22-columnist/26531; [accessed 2014-06-19] [WebCite Cache ID 6QS2WlxO4]

27. Capurro D, Cole K, Echavarría MI, Joe J, Neogi T, Turner AM. The use of social networking sites for public health practice and research: a systematic review. J Med Internet Res 2014;16(3):e79 [FREE Full text] [doi: 10.2196/jmir.2679] [Medline: 24642014]

28. Neiger BL, Thackeray R, Burton SH, Thackeray CR, Reese JH. Use of twitter among local health departments: an analysis of information sharing, engagement, and action. J Med Internet Res 2013;15(8):e177 [FREE Full text] [doi: 10.2196/jmir.2775] [Medline: 23958635]

29. Centers for Disease Control and Prevention. Best practices for comprehensive tobacco control programs. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014. URL: http://www.cdc.gov/tobacco/stateandcommunity/best_practices/pdfs/2014/comprehensive.pdf [accessed 2014-06-19] [WebCite Cache ID 6QS2gMqcO]

30. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants. 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]

31. HealthyPeople.gov. 2014. Health communication and health information technology URL: http://www.healthypeople.gov/2020/topicsobjectives2020/overview.aspx?topicid=18 [accessed 2014-06-19] [WebCite Cache ID 6QS2HUJBO]

32. Phua J. Participating in health issue-specific social networking sites to quit smoking: how does online social interconnectedness influence smoking cessation self-efficacy? J Commun 2013;63(5):933-952. [doi: 10.1111/jcom.12054]

33. Kim HS, Lee S, Cappella JN, Vera L, Emery S. Content characteristics driving the diffusion of antismoking messages: implications for cancer prevention in the emerging public communication environment. J Natl Cancer Inst Monogr 2013 Dec;2013(47):182-187. [doi: 10.1093/jncimonographs/lgt018] [Medline: 24395989]

Abbreviations

TCP: tobacco control program

XSL•FO
RenderX

Original Paper

The Role of Facebook in Crush the Crave, a Mobile- and Social Media-Based Smoking Cessation Intervention: Qualitative Framework Analysis of Posts

Laura Louise Struik[1*], RN, MSN; Neill Bruce Baskerville[2,3*], MHA, PhD

[1]Faculty of Health and Social Development, School of Nursing, University of British Columbia's Okanagan Campus, Kelowna, BC, Canada

[2]Propel Centre for Population Health Impact, Faculty of Applied Health Sciences, University of Waterloo, Waterloo, ON, Canada

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

[*]all authors contributed equally

Corresponding Author:
Laura Louise Struik, RN, MSN
Faculty of Health and Social Development
School of Nursing
University of British Columbia's Okanagan Campus
ART 360
3333 University Way
Kelowna, BC, V1V1V7
Canada
Phone: 1 2508647879
Fax: 1 2507647879
Email: laurastruik134@hotmail.com

Abstract

Background: Social networking sites, particularly Facebook, are increasingly included in contemporary smoking cessation interventions directed toward young adults. Little is known about the role of Facebook in smoking cessation interventions directed toward this age demographic.

Objective: The aim of this study was to characterize the content of posts on the Facebook page of Crush the Crave, an evidence-informed smoking cessation intervention directed toward young adults aged 19 to 29 years.

Methods: Crush the Crave Facebook posts between October 10, 2012 and June 12, 2013 were collected for analysis, representing page activity during the pilot phase of Crush the Crave. Of the 399 posts included for analysis, 121 were original posts, whereas the remaining 278 were reply posts. Posts were coded according to themes using framework analysis.

Results: We found that the original Crush the Crave Facebook posts served two main purposes: to support smoking cessation and to market Crush the Crave. Most of the original posts (86/121, 71.1%) conveyed support of smoking cessation through the following 7 subthemes: encouraging cessation, group stimulation, management of cravings, promoting social support, denormalizing smoking, providing health information, and exposing tobacco industry tactics. The remaining original posts (35/121, 28.9%) aimed to market Crush the Crave through 2 subthemes: Crush the Crave promotion and iPhone 5 contest promotion. Most of the reply posts (214/278, 77.0%) were in response to the supporting smoking cessation posts and the remaining 64 (23.0%) were in response to the marketing Crush the Crave posts. The most common response to both the supporting smoking cessation and marketing Crush the Crave posts was user engagement with the images associated with each post at 40.2% (86/214) and 45% (29/64), respectively. The second most common response consisted of users sharing their smoking-related experiences. More users shared their smoking-related experiences in response to the supporting smoking cessation posts (81/214, 37.9%) compared to the marketing Crush the Crave posts (11/64, 17%). With the exception of 4 posts, a moderator posted all the original posts. In addition, although 56.00% (18,937/33,815) of Crush the Crave Facebook page users were men, only 19.8% (55/278) of the reply posts were made by men. Finally, men were found to be more likely to express sarcasm or make strong assertions about quitting smoking and Crush the Crave than women.

Conclusions: The CTC Facebook page presents as a unique platform for supporting young adult smoking cessation at all stages of the cessation process. The findings of this study indicate that social networking sites, especially Facebook, warrant inclusion

XSL·FO
RenderX

in tobacco control efforts directed towards young adults. Research on effectiveness of the Facebook page for quitting smoking is needed.

KEYWORDS

qualitative research; young adult; smoking cessation; Internet; social media

Introduction

Background

Facebook has been recently included in the design and development of smoking cessation initiatives directed toward young adults. Because young adults (ages 18 to 29) are extensive users of Facebook [1], and because they represent the largest population of smokers in both Canada (21%) [2] and the United States (19%) [3], Facebook is increasingly included in contemporary smoking cessation initiatives directed toward this population. For example, Crush the Crave [4], a research-based smoking cessation app developed for young adults, has been integrated into Facebook, among other social media. The personal connections that can be formed through social media can be viewed as a form of social support and social support is effective in helping people quit smoking according to evidence gathered by the US Public Health Service expert panel [5]. However, despite the rapid growth in use of social media, research is in the early stages regarding how online social networks and opportunities for social support might or might not affect smoking cessation [6].

Social networking sites, particularly Facebook, remain under examined media for health promotion. Given the widespread uptake of Facebook among young adults, this medium warrants further investigation for its role in health behavior interventions directed toward this population, particularly in relation to smoking cessation. Therefore, the purpose of this study was to use framework analysis to characterize the content of the Crush the Crave Facebook posts and the associated responses.

Crush the Crave

In early 2012, a team of population health researchers, social media experts, and computer programmers developed and promoted Crush the Crave as an evidence-informed smoking cessation smartphone and social media app designed to help close the gap between existing smartphone apps [7] and evidence on what works in getting young smokers to quit smoking [8]. A panel of 5 experts in social media and tobacco cessation, a comparative analysis of the top 5 downloaded cessation apps, and 2 rounds of focus groups with young adult smokers were used to create the content and test the usability, design, and functionality of Crush the Crave.

Crush the Crave is available for Android and iOS devices in both English and French. Incorporating principles of persuasive technology for behavior change [9], Crush the Crave offers such features as a customized quit plan, the tracking of cravings and smoking habits, notifications of money saved and health improvements achieved, direct dial-up to telephone-based support, virtual awards that credit performance toward reaching milestones, evidence-informed credible information (eg, nicotine replacement therapy), and the ability to connect with a community of people for social support via Facebook.

Crush the Crave Facebook Page

The Crush the Crave Facebook page (Figure 1) is integrated within the smartphone app or can be accessed on its own through a browser. It is moderated by a social media expert and a small team with expertise in tobacco control. Individuals principally come to the Facebook page via the Google search engine or the Crush the Crave app. During the period of April 2012 to April 2013, Crush the Crave was piloted and promoted through Google and Facebook ads. Since this pilot phase, there has not been any active promotion of Crush the Crave.

As of November 19, 2013, the Crush the Crave Facebook page had 34,690 likes and a total reach of 7282 people. In all, 56.00% (4078/7282) of Crush the Crave Facebook users were men and 44.00% (3024/7282) were women. Most people reached (4369/7282, 60.00%) were between the ages of 18 and 34 years (the intended target group) and 57.00% (4151/7282) of Crush the Crave fans were from Canada. Posts to the Facebook page took place almost daily and usually included a photo and caption about quitting smoking. The total number of people who clicked on, liked, commented on, or shared Crush the Crave averaged 4000.4 (SD 4306.10) per week. User engagement in terms of likes, comments, and shares averaged 70.1 (SD 62.7) per post.

Figure 1. The Crush the Crave Facebook page.

Methods

Study Sample and Data Collection

For this study, 399 Crush the Crave Facebook posts, spanning from October 10, 2012 to June 12, 2013, were collected for analysis and entered into an NVivo version 10 qualitative software database (QSR International Pty Ltd, Burlington, MA, USA). This page activity represented the pilot phase of Crush the Crave. Posts were collected in reverse chronological order so that the most recent activity on the Crush the Crave Facebook page was represented. Sampling was driven by saturation of themes, where posts were collected until no new themes or subthemes were identified. Of the 399 posts, 121 were original posts, whereas the remaining 278 were responses to these posts. No direct contact with participants was sought for this study; data collection occurred solely through the gathering of posts freely available on the Crush the Crave Facebook page.

Data Analysis

The framework approach [10] was used to qualitatively analyze the Crush the Crave Facebook posts. Central to framework analysis is a series of interconnected stages (familiarization, identifying a thematic framework, indexing, charting, and mapping and interpretation) that enables the researcher to move back and forth between the data until a coherent account of the phenomenon is developed [11]. This analytic approach facilitated refinement of themes by both authors while maintaining a clear audit trail. Familiarization was achieved in this study through constant comparison [12] of posts in which coding categories were developed and entered in the NVivo

database. The first author coded all posts. To validate coding, both authors independently coded the first 51 posts and then compared for consistency. Any discrepancies in coding were discussed and resolved. In this way, each author was able to critically challenge one another on differing perspectives and any potential biases.

After the first 51 posts were coded, a thematic framework was developed by generating major themes and subthemes in relation to the original posts and categorizing the associated responses. This process was facilitated by asking iterative, analytic questions (eg, What is going on here? How are these different or the same?). Original posts were grouped according to their identified purpose (eg, supporting smoking cessation) and the responses were categorized according to their identified type (eg, sharing smoking-related experiences). To maintain the context of the responses to the original posts, they were listed under the messages from which they were derived and then categorized separately as a type of response. Throughout the coding process, regular meetings were held between the 2 authors to discuss and refine the thematic framework. Indexing was accomplished by coding each post in NVivo, with reliability checked by the second author through review of the NVivo file. The constant comparison method [12] was employed throughout the coding process to ensure internal homogeneity of the coding categories [13], whereby each post was coded into 1 category.

The fourth stage of framework analysis, charting, involved arranging summaries of the original posts and the associated responses in a table, which were grouped according to the identified themes and subthemes in relation to the purpose of

the original posts. This stage enabled the researchers to identify the range of data included in the original posts and to become familiar with the range of responses to these messages. The thematic framework was refined and two tables of data were developed: one that encompassed the purpose of the original posts according to each identified theme and subtheme and one that encompassed the various responses to these posts, which were grounded in the context of the themes and subthemes of the original posts. The final stage, mapping and interpretation, enabled the researchers to compare and contrast the original posts and responses while searching for patterns. At this stage, the original posts were grouped according to the finalized themes and subthemes, and responses were grouped together. Representative quotes were selected from the original posts and responses to illustrate key themes and subthemes.

Results

Characterizing the Original Crush the Crave Facebook Posts

Overview

We found that the original posts served two main purposes: to support smoking cessation and market Crush the Crave. Most posts (86/121, 71.1%) related to the major theme supporting smoking cessation, which included 7 subthemes (Table 1), and 35 of 121 posts (28.9%) related to the major theme marketing Crush the Crave, which included 2 subthemes (Table 2). The original Crush the Crave Facebook posts were primarily initiated by a Crush the Crave moderator. Only 4 of the posts were user-initiated.

Table 1. Posts supporting smoking cessation (N=86).

Subtheme	n (%)	Examples of post content	
Encourage cessation	34 (39)	Quitting is tough...but you're tougher ;)	
		Think of all the places a tobacco-free road can take you.	
Group stimulation	13 (15)	Fill in the blank! It's been ___days since my last cigarette.	
		Cool piece from '66. What would you rather spend your money on?	
Promote social support	12 (14)	You gotta love your quit buddy!	
		Good friends make quitting easier!	
Management of cravings	11 (13)	We found your next craving distraction-playing frisbee with a friend!	
		Next time you think about smoking, go to a mirror, take a breath, and make this face. Just try it. You'll see.	
Denormalize smoking	7 (8)	Hit the like button if you agree that smoking isn't sexy.	
		Fact: Not smoking makes you more kissable.	
Provide health information	5 (6)	Get outdoors and get active! Researchers found that physically active men were 36% more likely to have tried to quit smoking within the past year, while physically active women were 37% more likely to do so than nonactive women.	
		As soon as you stop smoking, you start repairing. (Thanks Australia!)	
Expose tobacco industry tactics	4 (5)	How do you feel about being labeled "a replacement"?	"The truth about tobacco" video [14]
		Why are young adults so important to tobacco companies?	

XSL•FO
RenderX

Table 2. Posts marketing Crush the Crave (N=35).

Subtheme	n (%)	Examples of post content	
Crush the Crave promotion	26 (74)	The new Crush the Crave is going to be Facebook-friendly!	
		Strut your way to a smoke-free lifestyle with Crush the Crave!	
iPhone 5 contest promotion	9 (26)	Did someone say iPhone 5? Our contest ends on June 15th!	
		Have you entered the Crush the Crave contest yet? You could WIN an iPhone 5!	

Supporting Smoking Cessation

The encourage cessation subtheme was present in many of the supporting smoking cessation posts (34/86, 39%). Posts within this subtheme contained positively framed messages offering encouragement to those trying to quit smoking, such as "Stay positive! You can do this!"

The group stimulation subtheme represented 13 of 86 (15%) posts under the major theme supporting smoking cessation. Generating interaction among the Crush the Crave Facebook users was the primary goal of these posts. Therefore, the posts consisted of questions and posting thought-provoking articles or images related to smoking. For example, a post contained a link to an article on new research demonstrating a link between cigarette smoking and damage to memory, learning, and reasoning, and asked users "Think there's truth to this?"

Management of cravings was another common subtheme (12/86, 14%). Posts under this subtheme included suggestions about ways to manage cravings (eg, "Find a peaceful place when cravings stress you out") and also invited users to reflect on their cravings (eg, "Quit Smoking Pro Tip: Reflect on your triggers to understand where your habit comes from").

The supporting smoking cessation posts also entailed a focus on promoting the act of social support for those trying to quit (11/86, 13%). The positive role that providing and receiving social support has on smoking cessation was emphasized in these posts. For instance, one post read, "Find a quit buddy, backup goes a long way."

Denormalize smoking was another subtheme under supporting smoking cessation and represented 7 of 86 posts (8%). These posts covered the positive social outcomes of being smoke-free. These positive outcomes included smelling good and looking attractive, such as "Did you know? Not smelling like cigarette smoke=more hugs."

Posts that provided health information made up 5 of 86 (6%) supporting smoking cessation posts. These posts provided facts about addiction, research on effective strategies for quitting smoking, the positive health effects of quitting, and information about the damaging effects of smoking on the body. For example, one post read, "Do you have a cigarette first thing in the morning? This recent study shows that smoking as soon as you wake up is even more unhealthy."

Expose tobacco industry tactics was another subtheme under supporting smoking cessation (4/86, 5%). The purpose of these posts was to invite Crush the Crave Facebook users to think about the recruitment strategies used by tobacco companies to increase their revenue. This is demonstrated in the post: "Class is in session! How does this Newport ad try to sell cigarettes to men?"

Marketing Crush the Crave

The Crush the Crave promotion subtheme represented most of the marketing Crush the Crave posts (26/35, 74%). The purpose of these posts was to promote the various mobile and social media platforms to access Crush the Crave (eg, "Did you know that Crush the Crave is on Google Play?"), the different features of the Crush the Crave app (eg, "Watching a funny video can help distract the crave! It's one of the features in our Quit Help section in the Crush the Crave app"), as well as news related to Crush the Crave (eg, "We got some radio play for new version of Crush the Crave today!").

The iPhone 5 contest promotion subtheme made up the remaining 9 of 35 (26%) marketing Crush the Crave posts. The focus of these posts was to motivate people to join Crush the Crave and download the Crush the Crave app. For instance, one post read, "Looking for some motivation? Enter our contest and you could win an iPhone 5!"

Characterizing the Crush the Crave User Reply Posts

Overview

Of the 278 reply posts, 214 (77.0%) were in response to the supporting smoking cessation posts, whereas the remaining 64 (23.0%) were in response to the marketing Crush the Crave posts. Women posted 218 (78.4%) of the reply posts, men posted 55 (19.8%), and 5 (1.8%) were posted by users that did not indicate their gender.

Supporting Smoking Cessation Reply Posts

The most common response (86/214, 40.2%) to the posts under the supporting smoking cessation major theme consisted of user engagement with the images associated with the original posts. Users primarily expressed their enjoyment of the images. For example, in response to an image of kittens hugging associated with a post promoting social support of cessation, a user responded: "Aw, so sweet!;))" A few users expressed that the images associated with the posts spurred on their smoke-free efforts. For example, in response to an image of puppies

associated with the caption "If I just hold your mouth shut, like this, you can't smoke" a user responded, "AH-adorable with a good reminder or push in the right direction." Conversely, some users indicated that although the images were appreciated, they were not enough to motivate their smoke-free efforts. For example, a user responded, "I love this cat...but nope, won't help me to stop smoking" to an image of a cat with the caption, "Maybe I can persuade you not to light that cigarette."

The second most common response (81/214, 37.8%) to the supporting smoking cessation posts consisted of users sharing their smoking-related experiences at various stages of the cessation process. These responses often consisted of users sharing their quit successes (eg, "Day 200 for me no smoking :3"). Some users would respond in detail and elaborate on their lived experiences with quitting smoking:

> I quit smoking 31 years ago. My husband had had a heart attack, we had to get someone to wash all the house and carpets and upholstery so the smoke smell would be gone. Because I loved him so much, and I wanted him to live with me much longer, he was my reason to stop. In order to be able to stop, you need a reason to help you along. Not so hard that way. And so worth it. Good luck all.

Crush the Crave Facebook users also shared their struggles with cessation, which was demonstrated in the following post: "I quit so many times...started again and again...last time I just made up my mind...threw them away and never looked back. It was hard but I've been smoke-free for 12 years now and very happy I did." Sharing personal management of cravings was another popular response in which users would describe what they found helpful in remaining smoke-free:

> I used an aide called Champix and I never looked back and never will!!!! And thanks, I am very proud of myself. I was a very heavy smoker. I smoked 25 to 30 a day and rolled my own. So harsh lol. Good luck to whomever is quitting. IT IS WORTH IT!!!!

Some Crush the Crave Facebook users also shared their smoking triggers. For example, one user described how she changed her daily routine so that she could quit smoking:

> For me, it was after a meal with my coffee. So I switched to tea and it made things easier. Or went without it and went for a walk because I never smoked outside. It has been 32 years now. I am so proud of myself.

Although seldom, some users would assert their current smoking status (eg, "Sorry, I do smoke"). Still others would express their intentions to quit (eg, "Well, Jan 1, 2013 will be the first day for me :) Wish me luck!").

Crush the Crave Facebook users also discussed tobacco control measures (11/214, 5.1%) in response to the supporting smoking cessation posts. For example, in response to a post asking users if they thought health warnings on cigarette packages should be mandatory, a user responded with the statement: "I don't think they should have to [put health warnings on cigarette packages]...you don't see any warning labels on liquor bottles or beer bottles or cans!! Drunken drivers kill..." Some users

expressed frustration with tobacco control efforts and suggested that a complete ban on cigarettes was needed: "...take THEM OFF THE shelf then, why can they keep selling something that kills you, when they pull a product off the shelf for WAY less health issues..."

Sometimes Crush the Crave Facebook users would express a negative attitude toward smoking (9/214, 4.2%) when engaging with the supporting smoking cessation posts. For example, a user described the social benefits of not smelling like cigarette smoke in response to a post that aimed to denormalize smoking: "More dates, more lov'n, more everything! It's not cool to smell like smoke!"

In all, 4 of 214 responses (1.9%) consisted of tags, which are links created by users to another person's Facebook timeline. The people who are tagged then receive a notification and are directed to the post that they were tagged in, facilitating their exposure to Crush the Crave content. Only 2 of 214 (0.9%) reply posts consisted of users sharing smoking-related facts, such as "It takes 7 years to recover." And 3 of 214 posts (1.4%) consisted of users encouraging the smoking cessation efforts of their peers: "If I can do it, anyone can. YOU CAN!!" In addition, 3 of 214 posts (1.4%) consisted of users expressing sarcasm. For example, in a post asking users how many days it's been since their last cigarette, a user responded with "too long." The remaining reply posts consisted of users expressing their misunderstanding of the posts (4/214, 1.9%) (eg, "What does the aurora borealis have to do with not smoking?"), a user sharing holiday wishes (1/214, 0.5%) (eg,"Happy Thanksgiving!"), and users posting content unrelated to the original post (12/214, 5.6%) (eg, "Do you ever have those dreams, where a bad person is chasing you and you're running in slow motion?").

Marketing Crush the Crave Reply Posts

Many (29/64, 45%) responses to the posts under the marketing Crush the Crave major theme consisted of users engaging with the images associated with the posts. A couple of images brought forward critiques by some of the Crush the Crave Facebook users. For example, the inclusion of an image portraying cigarettes that spelled out the word "quit" in a post promoting the Crush the Crave app received some criticism. One user made the following critique of the image: "It would make more sense if they [cigarettes in photo] were broken." Another user expressed that the use of images depicting cigarettes did not motivate their cessation efforts and, in fact, had the opposite effect: "This makes me want to smoke one." Of the 64 reply posts in response to the marketing Crush the Crave posts, 11 (17%) included users sharing their smoking-related experiences. Tagging others to the marketing Crush the Crave posts was another method of responding to the posts (6/64, 9%), indicating that users were promoting Crush the Crave to their friends.

Of the 64 replies to the marketing Crush the Crave posts, 6 (9%) consisted of sarcastic remarks about quitting smoking and what will help people quit. For example, in response to a post promoting Crush the Crave to help people quit smoking, a user posted, "Maybe if they put Justin Bieber on each pack of smokes then people will maybe quit smoking lol." Of the 64 reply posts, 4 (6%) consisted of users expressing a negative attitude toward

Crush the Crave, which is demonstrated in the following post: "Pathetic if you need an app to help you quit smoking." With the exception of one, the reply posts consisting of sarcasm or a negative attitude toward Crush the Crave were made by men.

In addition, 3 of 64 (5%) reply posts consisted of users gathering information about Crush the Crave. For example, a user wanted to know more about the availability of Crush the Crave on mobile phones: "Do you have a Blackberry version? Because it would be a shame if a Canadian-made app didn't work on a Canadian phone...:)." The remaining 3 of 64 (5%) reply posts consisted of peer encouragement of cessation (eg, "Quit smoking!! :)"), expression of a positive attitude toward Crush the Crave (eg, "Thumbs up"), and a comment unrelated to the Crush the Crave Facebook posts (eg, "We do a free package!").

Secondary Reply Posts

Of the 278 reply posts, 50 (18.0%) were secondary responses provided by moderators and peers. These secondary responses primarily consisted of the moderator offering clarification when there was misunderstanding, providing redirection when user responses were not related to the topic of the posts, and providing positive reinforcement for users sharing smoking-related experiences. For example, in response to a user sharing their quit success on the Crush the Crave Facebook page, a moderator responded, "This is a wonderful and inspiring story! Thank you so much for sharing it with us." Some responses to the original posts would prompt an ongoing conversation between the moderator and the users, or between the users, as shown in Textboxes 1 and 2.

Textbox 1. Sample conversation between the moderator and the users.

Post: Fill in the blank! I've pushed the CRAVE button _____ times today.

User #1 response: 3.

User #2 response: 0.

User #3 response: 9.

Moderator (asking the users who reported pushing the CRAVE button): Did you Distract the Crave with one of the options in the app?

User 1 response: Watermelon [distraction aid].

Textbox 2. Sample conversation between the users.

Although seldom, some users would engage with the responses of their peers, such as this conversation:

Post: Does anyone remember the *Friends* (TV show) episode when Chandler tries to quit smoking? Maybe he should have tried our app!

User #1 response: 3 years plus now nonsmoking thanks to the help of Champix.

User #2 response: Posted article about Champix being linked to suicides.

User #1 response: The article forgot to mention that using tobacco kills 40,000 Canadians each year. I'd take my chances with Champix.

Discussion

Principal Findings

This study addresses gaps in the literature by investigating the role of Facebook as part of a smoking cessation intervention directed toward young adults. The findings of this study reveal that social networking sites, such as Facebook, can be harnessed for supporting young adults who are trying to quit smoking or have become smoke-free.

The primary purpose of the Crush the Crave Facebook page was to support smoking cessation. The supportive nature of the Crush the Crave Facebook page was first cultivated through the original posts, and then maintained through the reply posts, as well as through the secondary responses provided by the Crush the Crave Facebook page moderators, where positive reinforcement, clarification, and redirection was provided in response to user posts. This finding extends previous research [15] and indicates that Facebook is not only a context that can host supportive exchanges, but also that the original posts on a Facebook page may play a major role in establishing a supportive context for these exchanges. Once established through the original posts, the supportive nature of the Crush the Crave Facebook page was then maintained and continued through the responses made by a variety of sources (eg, moderator, smoke-free individuals, individuals trying to quit) in a variety of ways (eg, offering information, providing positive reinforcement, sharing personal experiences). Given that social support increases the likelihood of quitting and success in remaining smoke-free [16,17], the supportive context established on the Crush the Crave Facebook page is encouraging.

Leadership on the Crush the Crave Facebook page was primarily provided by the Crush the Crave moderators. The moderators posted almost all the original posts and continually sought to encourage individuals in their cessation efforts, promote engagement by posing questions or topics of interest to the group, and moderate discussions. Given that individuals on the page were not likely to initiate engagement on the page, it appears that users relied on the direction of a moderator. Although leadership on the page may evolve over time, this finding highlights the key role that moderators play in stimulating group activity and in directing and tailoring page content to the group. Other researchers who have investigated the role of social forums in health promotion interventions report similar findings [15,18]. For example, Ploderer and colleagues [15] found that users of a Facebook page as part of a smoking cessation intervention also relied on the direction of a moderator

rather than on the contributions of its members. It has been suggested that moderators play a pivotal role in influencing health behavior change and that the absence of a moderator may discourage users and hinder their motivation and self-confidence in changing their health behaviors [18,19]. Given increasing evidence that a moderator plays a pivotal role in harnessing the benefits of including social networking sites in health promotion interventions, consideration for ongoing inclusion of moderators for providing support and guiding page content is warranted.

In relation to the reply posts, we found that individuals on the Crush the Crave Facebook page primarily shared their smoking-related experiences. The sharing of personal experiences on the Crush the Crave Facebook page is similar to the findings of a recent study by Brandt and colleagues [20], demonstrating that Facebook is used as a platform for self-reflection and evaluation of one's cessation efforts [20]. This finding adds to increasing evidence that social networking sites, particularly Facebook, may play a critical role in helping individuals reach their cessation goals.

We found that Crush the Crave Facebook users often posted about their various tobacco use behaviors and experiences, from their current smoking status (eg, "Sorry, I do smoke") to their success in becoming smoke-free (eg, "I quit smoking just over 1 month ago and I've never felt better"). Although infrequent, members on the page would share their cessation challenges. The display of quit struggles and smoking relapses on the Crush the Crave Facebook page is similar to the findings of Ploderer and colleagues [15], who also found that members of a smoking cessation Facebook page were willing to share their quit struggles. This finding adds to increasing evidence that Facebook pages dedicated to supporting a particular health behavior, such as the Crush the Crave Facebook page, are trusted sources of social support. In light of evidence that individuals are not likely to share their health behavior struggles on their personal Facebook pages because of concern for embarrassment [21,22], the inclusion of Facebook pages as part of health behavior interventions provides individuals with a support network that they can tap into and share a broad range of experiences.

More frequently, users would share cessation success stories. This finding is encouraging because successful quitters are a valuable resource for others who are trying to quit, and also provide normative influence within a social network [17,23]. In addition, abstinent smokers in a social network are reminded of their cessation journey, which strengthens their wish to remain smoke-free [23]. Furthermore, individuals who are trying to quit can draw motivation and encouragement from reading the stories of those who have succeeded in their cessation [23], thereby adding to the supportive nature of the Crush the Crave Facebook page.

Gender differences in user engagement on the Crush the Crave Facebook page were highlighted in the findings. Consistent with recent reports indicating that women are more likely than men to engage on social networking sites, such as Facebook [24], we found that most responses to both the original posts and other user posts were made by women. Although men were found to be less active, they also represented the largest users

of the Crush the Crave Facebook page, indicating that the Crush the Crave Facebook page may be a promising avenue for reaching young men who smoke. In light of evidence that young adults are less likely to participate in smoking cessation programs [25,26], this finding highlights the potential of engaging young adults in smoking cessation interventions via Facebook, especially young women.

Gender differences in the communication styles of men and women were also noted on the Crush the Crave Facebook page. Women's communication on the Crush the Crave Facebook page emphasized supportiveness and connection with others, whereas men were more likely to express sarcasm and make strong assertions about Crush the Crave or quitting smoking. This finding is consistent with the findings of previous studies investigating communication on Facebook specifically [21], and social media more broadly [27]. Although we must be cautious about falling into heteronormative assumptions when approaching smoking cessation online, consideration for the stereotypical communication styles of men and women displayed on Facebook is warranted.

Limitations and Future Research

The findings of this study need to be considered in light of several limitations. We were limited in describing Crush the Crave Facebook users through Facebook Insights. In addition, gender and age group were the only demographics freely available on the Crush the Crave Facebook page for the sample of users included in this study. Because this study specifically examined Facebook, the findings may not be applicable to other social networking platforms. Finally, it is important to be mindful of the continuously evolving nature of the Crush the Crave Facebook page over time (eg, membership, introduction of new topics, new group leaders, new moderators, new promotion campaigns) and that these findings are situated in a particular timeframe. In particular, the Crush the Crave Facebook page activity for this study was representative of the pilot phase of Crush the Crave to assess feasibility. Work is underway to launch a randomized controlled trial of Crush the Crave to determine its effectiveness in quitting smoking

Future research on the effectiveness of mHealth and social media interventions in helping people quit smoking is needed. Research on how social support moderates the relationship between use of social media and quitting smoking is also needed. In addition, given the variability in smoking rates across population subgroups, understanding what subpopulations of smokers can benefit the most from social media and mobile-based smoking cessation interventions is needed. Furthermore, given the noted gender differences on the Crush the Crave Facebook page, further research is needed investigating participation of men and women on the Facebook page and how their smoking behaviors are influenced through the page. Once demonstrated as cost-effective and for whom, policy makers can determine how best to integrate mHealth and social media interventions within comprehensive tobacco cessation systems [28,29].

Conclusions

Social networking sites are increasingly included in contemporary smoking cessation interventions. Given the

relative lack of evidence in relation to online smoking cessation initiatives, this exploratory analysis contributes to knowledge about the ways in which social networking sites, such as Facebook, fit within the overall cessation picture. We found that the Crush the Crave Facebook page is a useful tool for providing individuals with opportunities to give and receive support throughout the smoking cessation process, thereby adding a new and innovative dimension to smoking cessation interventions directed toward young men and women.

Authors' Contributions

LS and NB conceived the study. NB was senior scientist and oversaw the study. NB developed the intervention in collaboration with IMP Canada, DH, and SF. LS oversaw the conduct of the analysis for the paper, preparation of the manuscript, and NB contributed. LS and NB interpreted the findings. All authors contributed to the final manuscript. All authors had full access to all data. NB is guarantor for the paper.

Conflicts of Interest

None declared.

References

1. Duggan M, Brenner J. The demographics of social media users-2013. Washington, DC: Pew Internet & American Life Project; 2012. URL: http://pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx [accessed 2013-12-09] [WebCite Cache ID 6LkOCjRIN]

2. Health Canada. 2012. Canadian tobacco use monitoring survey (CTUMS) URL: http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/research-recherche/stat/_ctums-esutc_2012/ann_summary-sommaire-eng.php [accessed 2013-12-09] [WebCite Cache ID 6LkOGztHV]

3. Centers for Disease Control and Prevention. 2013. Adult cigarette smoking in the United States: Current estimate URL: http://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/ [accessed 2013-12-09] [WebCite Cache ID 6LkOKN4ar]

4. Crush the Crave. URL: http://www.crushthecrave.ca/ [accessed 2014-06-23] [WebCite Cache ID 6QXt0t9xy]

5. Clinical Practice Guideline Treating Tobacco UseDependence 2008 Update Panel, Liaisons,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. [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]

6. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants. 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]

7. Abroms LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. Am J Prev Med 2011 Mar;40(3):279-285 [FREE Full text] [doi: 10.1016/j.amepre.2010.10.032] [Medline: 21335258]

8. Fiore M. Treating Tobacco Use and Dependence: 2008 Update: Clinical Practice Guideline. Rockville, MD: DIANE Publishing Company; 2008.

9. Fogg BJ, Hreha J. A method for matching target behaviors with solutions. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. Pursuasive Technology. Berlin: Springer; 2010:117-131.

10. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RH, editors. Analyzing Qualitative Data. London: Routledge; 1994.

11. Ritchie J, Lewis J. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage Publications; 2003.

12. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, IL: Aldine Pub Co; 1967.

13. Lincoln YS, Guba EG. Naturalistic Inquiry. Beverly Hills, CA: Sage Publications; 1985.

14. YouTube. The Truth About Tobacco: How Much is a Life Worth? URL: http://www.youtube.com/watch?v=2m7-zIa6-Es [accessed 2014-06-25] [WebCite Cache ID 6QaHjrVIN]

15. Ploderer B, Smith W, Howard S. Patterns of support in an online community for smoking cessation. New York, NY: ACM; 2013 Presented at: 6th International Conference on Communities and Technologies; June 29-July 2, 2013; Munich, Germany. [doi: 10.1145/2482991.2482992]

XSL·FO
RenderX

16. Cohen S. Social relationships and health. Am Psychol 2004 Nov;59(8):676-684. [doi: 10.1037/0003-066X.59.8.676] [Medline: 15554821]

17. Westmaas JL, Bontemps-Jones J, Bauer JE. Social support in smoking cessation: reconciling theory and evidence. Nicotine Tob Res 2010 Jul;12(7):695-707. [doi: 10.1093/ntr/ntq077] [Medline: 20513695]

18. Lindsay S, Smith S, Bellaby P, Baker R. The health impact of an online heart disease support group: a comparison of moderated versus unmoderated support. Health Educ Res 2009 Aug;24(4):646-654 [FREE Full text] [doi: 10.1093/her/cyp001] [Medline: 19251770]

19. Syred J, Naidoo C, Woodhall SC, Baraitser P. Would you tell everyone this? Facebook conversations as health promotion interventions. J Med Internet Res 2014;16(4):e108 [FREE Full text] [doi: 10.2196/jmir.3231] [Medline: 24727742]

20. Brandt CL, Dalum P, Skov-Ettrup L, Tolstrup JS. "After all--it doesn't kill you to quit smoking": an explorative analysis of the blog in a smoking cessation intervention. Scand J Public Health 2013 Nov;41(7):655-661. [doi: 10.1177/1403494813489602] [Medline: 23696257]

21. Morris E. Digital Collections University of Colorado Boulder. 2013. She "likes" it, he doesn't: Gender differences in Facebook communication behaviors URL: http://digitool.library.colostate.edu/ //exlibris/dtl/d3_1/apache_media/L2V4bGlicmlzL2R0bC9kM18xL2FwYWNoZV9tZWRpYS8yMTEwNzU=.pdf [accessed 2014-06-26] [WebCite Cache ID 6Qd7SWoGI]

22. Newman MW, Lauterbach D, Munson SA. It's not that I don't have problems, I'm just not putting them on Facebook: Challenges and opportunities in using online social networks for health. : ACM; 2011 Presented at: ACM 2011 conference on Computer supported cooperative work; March 19-23, 2011; New York, NY p. 341-350. [doi: 10.1145/1958824.1958876]

23. Bartholomew LK, Parcel GS, Kok G. Planning Health Promotion Programs: An Intervention Mapping Approach. San Francisco, CA: Jossy-Bass; 2011.

24. Hampton KN, Goulet LS, Marlow C. Why most Facebook users get more than they give. Washington, DC: Pew Internet & American Life Project; 2012 Feb. URL: http://pewinternet.org/Reports/2012/Facebook-users.aspx [accessed 2013-12-09] [WebCite Cache ID 6LkORkKai]

25. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. Am J Public Health 2010 Jul;100(7):1282-1289 [FREE Full text] [doi: 10.2105/AJPH.2009.165449] [Medline: 20466971]

26. Minian N, Schwartz R, Di Sante E. Ontario Tobacco Research Unit. 2010. Impact of the smoking cessation system on young male smokers URL: http://otru.org/wp-content/uploads/2012/06/special_yms.pdf [accessed 2014-06-27] [WebCite Cache ID 6Qd86oX56]

27. Thomson R, Murachver T. Predicting gender from electronic discourse. Br J Soc Psychol 2001 Jun;40(Pt 2):193-208. [Medline: 11446227]

28. Institute of Medicine. Ending the Tobacco Problem: A Blueprint for the Nation. Washington, DC: National Academies Press; 2007.

29. Smoke-free Ontario Scientific Advisory Committee. Evidence to Guide Action: Comprehensive Tobacco Control in Ontario. 2010. URL: http://www.publichealthontario.ca/en/eRepository/ Evidence%20to%20Guide%20Action%20-%20CTC%20in%20Ontario%20SFO-SAC%202010E.PDF [accessed 2014-06-27] [WebCite Cache ID 6Qd8Ir1EX]

Review

Social Network Sites as a Mode to Collect Health Data: A Systematic Review

Fahdah Alshaikh[1], BSc (Hons), MPH; Farzan Ramzan[1], BSc (Hons), MSc; Salman Rawaf[1], MD, FRCP, FFPH, PhD; Azeem Majeed[1], MD, FRCP, FRCGP, FFPH

School of Public Health, Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Fahdah Alshaikh, BSc (Hons), MPH
School of Public Health
Department of Primary Care and Public Health
Imperial College London
324 The Reynolds Building
St Dunstan's Road
London, W6 8RP
United Kingdom
Phone: 44 (0)20 7594 7386
Fax: 44 (0)20 7594 0854
Email: fahdah.alshaikh09@imperial.ac.uk

Abstract

Background: To date, health research literature has focused on social network sites (SNS) either as tools to deliver health care, to study the effect of these networks on behavior, or to analyze Web health content. Less is known about the effectiveness of these sites as a method for collecting data for health research and the means to use such powerful tools in health research.

Objective: The objective of this study was to systematically review the available literature and explore the use of SNS as a mode of collecting data for health research. The review aims to answer four questions: Does health research employ SNS as method for collecting data? Is data quality affected by the mode of data collection? What types of participants were reached by SNS? What are the strengths and limitations of SNS?

Methods: The literature was reviewed systematically in March 2013 by searching the databases MEDLINE, Embase, and PsycINFO, using the Ovid and PubMed interface from 1996 to the third week of March 2013. The search results were examined by 2 reviewers, and exclusion, inclusion, and quality assessment were carried out based on a pre-set protocol.

Results: The inclusion criteria were met by 10 studies and results were analyzed descriptively to answer the review questions. There were four main results. (1) SNS have been used as a data collection tool by health researchers; all but 1 of the included studies were cross-sectional and quantitative. (2) Data quality indicators that were reported include response rate, cost, timeliness, missing data/completion rate, and validity. However, comparison was carried out only for response rate and cost as it was unclear how other reported indicators were measured. (3) The most targeted population were females and younger people. (4) All studies stated that SNS is an effective recruitment method but that it may introduce a sampling bias.

Conclusions: SNS has a role in health research, but we need to ascertain how to use it effectively without affecting the quality of research. The field of SNS is growing rapidly, and it is necessary to take advantage of the strengths of this tool and to avoid its limitations by effective research design. This review provides an important insight for scholars who plan to conduct research using SNS.

KEYWORDS

social network; social media; Internet; systematic review; young people; health education; health behaviors

XSL•FO
RenderX

Introduction

Overview

Since their introduction, social network sites (SNS) have attracted individuals, businesses, social organizations, and lately health organizations and providers. There are millions of users, each with a different purpose for using these networks.

The purpose of this review is to focus on those networks that are defined as "Web-based services that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system" [1].

Social network sites and social media include all types of online social platforms that allow participants to share interests and opinions and many other social interactions. The use of these platforms is becoming dominant among all Internet usage purposes, and today Web content often has the feature to share or link to SNS. It seems that the importance of a topic is linked to its presence in SNS [2].

Social networking is not just about being on a website. It comprises a community that shares and interacts. It is a powerful community that has shifted the concept of media and is rapidly and extensively penetrating society [3].

Social Network Sites in Health Research

For researchers, SNS is an environment where sharing information, knowledge, interest, and opinion is meaningful and fun, which makes it ideal for conducting research [2]. The promise that SNS held for health has been explored and discussed in previous publications. SNS may play a role for health in two ways: (1) the presence of health organizations on SNS makes them more approachable and accessible, and (2) SNS may be an effective way of helping patients with chronic diseases manage their health conditions. The importance of SNS is reflected by increasing efforts within health sectors and organizations to embrace SNS [3,4].

However, efforts towards using SNS are still in their infancy and more inventive interventions and other ways of benefiting from SNS are yet to be explored and discussed [3]. One of the many possible uses of SNS for health is using this powerful platform to collect data and recruit for research studies.

Potential of Social Network Sites for Data Collection

Interest in online social networks has been increasing over the past few years as a result of the huge adoption of this technology all over the world.

The literature shows that SNS has been used by researchers as a source of information about user characteristics, patterns of friending, and usage behavior [5]. These social networks have become a modern source for information and data gathering. They have evolved into a dynamic and accurate source of gathering information because they contain a feature not found in traditional media: active and two-way participation [6].

SNS are also extraordinary marketing tools, able to reach almost any type of person, which changes communication from "one-to-one" to "many-to-many" [7]. Finally, they have become sources of collecting timely information, converting data into profitable results at a faster rate. They contain great opportunities for future research in public health because they can be a great way to reach hidden and hard-to-reach groups [8]. Yet there is still relatively little direction on how SNS can be used in health research and whether they can provide valid and reliable data.

Research Objectives

The aim of this study was to systematically review the available literature and explore the use of SNS as a mode of collecting data for health research. The review aims to answer four questions: (1) Does health research employ SNS as a mode of collecting data? (2) Is data quality affected by the mode of data collection? (3) What types of participants were reached by SNS? (4) What are the strengths and limitations of SNS?

Methods

Systematic Review

The literature was reviewed systematically by searching bibliographic databases, MEDLINE, Embase, and PsycINFO, in March 2013 using the Ovid and PubMed interface for the period 1996 to the third week of March 2013, using the following keyword combinations: (Online Social Networks or Online Social Sites or Social Media) AND (Health). In addition, a manual search was undertaken, searching the reference list of all included studies.

The review was conducted by 2 reviewers independently. Search results were extracted to an Endnote database, inclusion and exclusion processes were recorded, and all abstracts and titles were reviewed. The initial selection criteria were (1) Intervention: The review focuses specifically on using SNS as a mode of collecting data, rather than as a social intervention (eg, support group), (2) Time and place: Studies produced at any time and place will be included in the search strategy, (3) Study participants: can include community or patients, and (4) Outcomes: included studies must contain outcomes related to the data collection mode, for example, response rate, completeness, missing data, timeliness, cost, and perception of privacy and anonymity. There were no language restrictions to ensure that as many studies as possible were assessed for relevance to the review.

Studies were excluded if they examined SNS participant interaction rather than SNS as a mode of collecting data, if they did not involve SNS, or if the article was a general discussion paper that did not present data or methods. All included studies had to specify the use of SNS as a tool to collect self-reported health data.

The review focused on the quality of data, the strengths and limitations of the mode, and reported strategies that facilitate the data collection. After inclusion based on title and abstract, full articles were retrieved and data extracted with a predesigned extraction form that included a checklist to assess the quality of each included study. The checklist was developed by the Centre for Reviews and Dissemination (CRD) [9]. Studies that
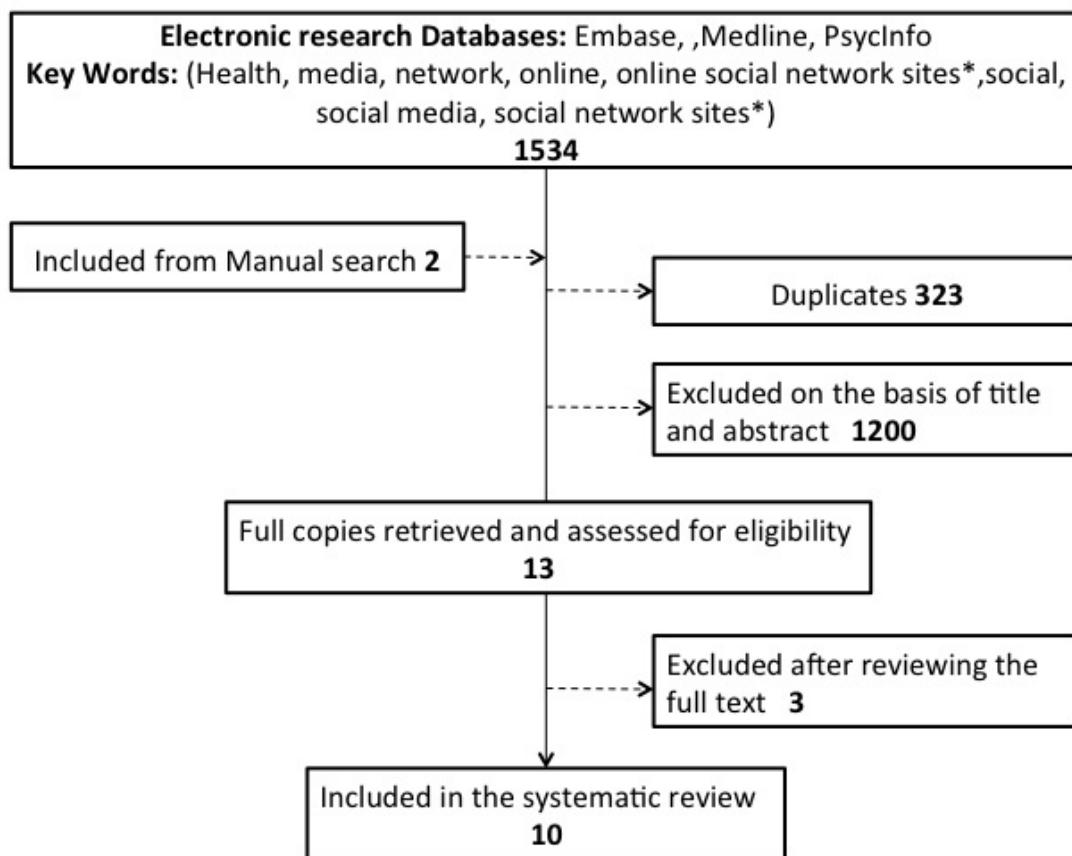
scored 5 out of 7 or below were considered as low quality, and above 5 were considered as high quality.

Data Analysis

The wide variety of methodologies and outcomes of included studies limited the possibility to carry out meta-analyses. For example, study populations were different and for the studies with similar populations, they had different methods. A descriptive qualitative analysis was carried out to answer the four research questions.

The PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) checklist was followed in this systematic review [10]. The checklist items are essential for transparent reporting of a systematic review, and the author covered most of these items except items related to meta-analysis, which was not applicable for this review (Multimedia Appendix 1).

Results

Overview

A total of 1534 citations were identified using the search strategy from the electronic databases and search results combined with articles identified by searching manually with duplicates removed; 1213 citation titles and abstracts were reviewed.

A full text assessment was undertaken on the 13 papers that appeared to meet the inclusion criteria based on title and abstract. Of these, 3 were excluded: 2 not employing SNS directly to collect data and 1 discussion paper with no results. A total of 10 papers were reviewed and assessed. Table 1 depicts the search results from each database, and Figure 1 illustrates the review process.

Figure 1. Summary of the systematic literature review process.

Included Studies

All included studies were cross-sectional and primarily used self-reported data (Multimedia Appendix 2). While the majority of studies undertook quantitative analysis, 1 study was based on qualitative focus group data. All included studies are defined as high quality after they were assessed by the CDR checklist (Multimedia Appendix 3).

Table 1. Database search results.

Database	Database provider	Years searched	Number of citations retrieved (n=1536)
MEDLINE	OVID	1996 to March week 1, 2013	403
PsycInfo	OVID	1996 to March week 2, 2013	142
Embase	OVID	1996 to March week 3 2013	609
MEDLINE	Pubmed	1996 to March week 3, 2013	380
Manual search	–		2

Does Health Research Use Social Network Sites as Mode of Collecting Data?

Overview

The large number of search results indicate that health research is using SNS in many forms, with the majority of studies investigating the effect of SNS use on health seeking behavior and knowledge. However, there are a limited number of studies that focus on SNS as a tool for research recruitment or data collection; only 10 studies of 1213 looked at this potential.

Type of Collected Data

Of the included studies, 9 collected survey data through SNS, while Levine (2011) conducted an online focus group within a SNS [11]. Fenner (2012) used SNS for recruitment, rather than only for data collection [12]. In summary, most of the included studies collected quantitative data, and only one, Levine (2011) collected qualitative data through MySpace and explored possibilities to increase response rates [11]. It is more convenient to collect quantitative data rather than qualitative through SNS, as the latter may require more resources.

Data Collection

As this review is looking at SNS as mode to collect health research data, all included studies reached their participants through SNS; however, different approaches were undertaken. Heather (2009) distributed an invitation for an online survey on 8 SNS related to pregnancy and baby health. This approach resulted in 288 valid surveys in a 2-month period [13].

Levine (2011) conducted a focus group by forming a MySpace profile embedded within a chat room and sending invitations to members to join their network; 738 people joined the study's network. Participant selection was based on reviewing SNS profiles, inviting specific participants that met the inclusion criteria [11]. Although this methodology produced an accurate set of qualitative data, it required several staff members to aid with the recruitment process.

Woolley (2012) was interested in monitoring the impact of a specific Facebook health fan page, "Get Up And Do Something (GUADS)", on participant health seeking actions and behavior. The study used the fan page to collect data by inviting all fans to participate in an online survey [14]. Although the study reported the use of multiple recruitment methods, no further explanation was given as to the recruitment process.

Of the included studies, 4 used Facebook ads, an online advertising affiliate program, which is a powerful targeted advertisement method. Fenner (2012) assessed the feasibility of recruiting young females using this method. In total, 278 participants were recruited, 139 chose to participate in the online survey, and the remaining 139 opted to physically attend the research center [12]. Ramo (2012), Lohse (2013), and Lord (2011) also used Facebook ads to advertise an online survey to their target populations [15-17].

In addition to Facebook ads, 2 other applications within Facebook were used by health researchers: Facebook Event (a new feature on Facebook where the plugin gives the fan page administrator the ability to add details about upcoming events, eg, Event Name, Location, and Date) and Facebook Poll (a page widget that gives the fan page administrator the ability to add "poll questions" with a voting system and closing date) [18,19]. Finally, Shindel (2012) investigated the association of high risk for sexual dysfunction of women who have sex with women (WSW). Individuals were invited to participate by emailing the entire member list of online social networks catering to WSW [18].

Is Data Quality Affected by the Mode of Data Collection?

Overview

In order to address this question, we looked at data quality indicators such as response rate, completion time, dropout rate, timeliness, missing data, and cost. A comparison was undertaken where appropriate, as studies differed substantially in methodology and population. However, none of the included studies looked at the quality of data in depth or performed any type of analysis regarding quality. Response rate, cost, timeliness, and missing data were reported in some studies.

Response Rate

Response rate is defined as the "Number of participants who completed a questionnaire" divided by the 'Total number of participants who were asked to participate" [20]. The highest response rate 27% (N=2583) was reported by Lord (2011). The survey had been advertised on Facebook for 2 weeks and targeted a young population with no strict inclusion criteria [17].

The lowest response rate, 2.2% (N=738) reported by Levine (2011), resulted from the recruitment method that was used. Participants were invited to a synchronous online focus group (during a specified time). Later changing to asynchronous (ie, no specified time) caused a slight increase to 7.2% (N=250) [11].

The collection of qualitative data is challenging because it requires more effort and staff time as reported by the study

authors; in some cases, calculating the response rate was not possible [14]. To summarize (Table 2), the reported response rates ranged between 2% to 27%, with an average of 12%.

Table 2. Reported response rates of included studies.

Reference	Study ID	Response rate, %	Participated/Reached
[17]	Lord, 2011	26.67	689/2583
[19]	Cucchetti, 2012	22.80	2414/10584
[16]	Lohse, 2013	17.42	18/465
[14]	Woolley, 2012	11.19	90/804
[15]	Ramo, 2012	10.69	1548/14808
[12]	Fenner, 2012	6.93	551/7940
[11]	Levine, 2011	2.17(synchronous)	16/738
		7.20 (asynchronous)	18/250

Cost and Timeliness

Of the studies, 4 used Facebook in their recruitment strategy; 3 were able to report on the cost and timeliness of data collection (Table 3) [12,15,16].

The criteria used by the authors in targeting specific participants will inadvertently affect the number of participant responses. Strasser (2012) set out to recruit 100 participants and closed the survey as soon as this target was met [21]. Evidently, time was not a key factor and the recruitment process would have continued as required.

However, cost is an implication that must be considered. Studies may have a limited budget and could recruit only as many participants as possible within a specific timeframe. Ramo (2012) reported the lowest cost per participant: $4.28 over a 13-month period yielding 1548 participants who were young smokers [15]. Lohse (2013) and Fenner (2012) targeted females within specific age ranges, which could account for the higher costs reported [12,16].

Nevertheless the highest reported cost $20.14 was considered favorable over the cost of traditional methods of recruiting as reported by Fenner (2012) [12].

Table 3. Reported cost and timeliness of included studies.

Reference	Study ID	Duration (days)	Participants	Response per day	Cost ($US)		
					Per participant	Per click	Total
[15]	Ramo, 2012	390	1548	4	4.28	0.45	6628.24
[16]	Lohse, 2013	19	62	3	9.26	1.28	596.71
[12]	Fenner, 2012	150	278	2	20.14	0.67	5598.92
[21]	Strasser, 2012	16	100	1.6	–	–	–

Other Quality Indicators

Fenner (2012) was the only author to report on missing data. For the demographic questions, this did not exceed 5%, and for remaining questions, this value was less than 8%. The author considered this a positive indicator on the quality of data [12].

Lohse (2013) reported that the completion rate of the survey was 93.5%, which is indicative of good data quality [16]. The validity of the data was reported by Lord (2011) as 76%; however, no further explanation was provided as to how this was assessed [16].

What Types of Participants Were Reached by Social Networking Sites?

Although SNS is a tool that can be widely used to recruit participants, it may be more effective for certain groups; for example, if targeting an aging population, one has to take into account that this group may not be as computer literate and therefore less likely to use SNS. The types of participant more suited to SNS (Table 4) would be a younger population and those that are "hard to reach", for example, a homosexual population.

Table 4. Types of participants targeted by SNS.

Reference	Study ID	Participant type	Target age, years
[13]	Heather, 2009	Pregnant women	>18
[12]	Fenner, 2012	Female	16-25
[16]	Lohse, 2013	Female	18-45
[18]	Shindel, 2012	WSW	>18
[21]	Strasser, 2012	MSM	No targeted age
[11]	Levine, 2011	Youth	16-24
[17]	Lord, 2011	Youth	18-25
[15]	Ramo, 2012	Young smokers	18-25
[14]	Woolley, 2012	Community	>18
[19]	Cucchetti, 2012	Community	No targeted age

What Are the Strengths and Limitations of Social Networking Sites?

One of the most reported strengths of SNS is that it is an effective recruitment method. This was stated in 4 studies, which successfully reached young age groups [15], females [12], low-income females [16], and MSM [21]. All these populations were defined by researchers as hard-to-reach groups. Facebook in particular was considered by Ramo (2012) as a successful mechanism to reach and recruit a young age group in smoking-related health research, which is normally a challenge [15]. In addition, Levine (2011), Fenner (2012), and Ramo (2012) reported that SNS proved to be much more cost effective over other traditional methods of recruiting in health research [11,12,15]. SNS can also provide representative and valid data. Fenner (2012) indicated that the SNS sample yielded demographically representative data, and Lord (2011) stated SNS provided a rich pool of qualitative and quantitative valid data [12,17].

Another strength of SNS is that using online focus groups can be an easy and simple process if conducted asynchronously through SNS, which allows one to capture the exact language of participants to analyze [11]. Finally, an important strength of SNS for health surveys and research is the potential for sharing and invitation within the network, enabling surveys to be diffused rapidly between SNS participants [19].

The predominant limitation of SNS for collecting data was that it may introduce self-selection bias, and when there is a self-selection bias usually there is a sample bias and representative and generalizability issues. Strasser (2012) has also stated that self-reported data may affect the reliability and validity of results [21].

Discussion

Principal Findings

This comprehensive review addressed our original research questions and found a gap in the literature for evaluating the effectiveness of SNS as a tool in health research. The findings demonstrate that SNS is considered a research tool that can reach wide audiences and simplify the data collection process for health research, especially quantitative data, along with a wide range distribution of surveys reaching many participants through SNS.

SNS is a powerful tool that can provide a wealth of information about research participants and has the potential to capture good quality data, as some of the included studies have shown. However, SNS self-reported data may introduce self-selection bias, sampling bias, or other generalizability/reliability issues. This aspect was not fully investigated in the included studies of this review, which therefore indicates the need for future research or systematic reviews to focus on these issues.

In this review, Facebook was used in 8 out of 10 of the included studies, which indicates its strong potential as a tool for conducting health research. Many features within Facebook empower the research process, for example, Facebook ads, polls, events, and insights. The potential of Facebook needs to be highlighted especially in health research where validity is of utmost importance for research results. Hence, further studies assessing its potential in health research are needed.

Strengths and Limitations

A number of strengths were highlighted in this review. First, it was an overview of the existence of SNS use in health research literature illustrating the strengths and limitations of this method in data collection. Second, it was a comprehensive and explicit review with broad inclusion criteria that led to a review of 1213 studies, highlighting the gap in the literature regarding the use of SNS as a tool and its effect on data quality.

A limitation of this review is the heterogeneity of the included studies. Although all used SNS to collect data, their individual objectives, populations, and outcomes are unique. Analyses were found to be primarily descriptive.

Systematic Review Outcomes

SNS can be suitable for health research and was claimed to be an effective tool to collect data, but more research is required to look more closely at its effectiveness as a tool. Comparative research that compares SNS with other data collection modes would be valuable in highlighting differences between the quality of data obtained, costs incurred, and samples obtained. This review indicates that the quality of collected data was not

assessed thoroughly; although for surveys and online questionnaires, it led to an acceptable level of validity. Yet, SNS use for data collection proved to be more successful when young age groups were targeted. Finally, Facebook SNS was used in a number of included studies in this review and highlighted that it is a powerful tool that provides multiple features that can be used to improve online health research.

Conclusions

This review concludes that SNS has a niche in health research, but we need to ascertain how to use it effectively without affecting the quality of research. The field of SNS is growing rapidly and researchers need to take advantage of the strengths of this tool and to avoid its limitations by employing effective research design.

Acknowledgments

Authors' Contributions

FA developed the concept for the research and the paper. FA and FR collected data and analysis. FA, SR, AM, and FR finalized the text.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA Checklist.

[PDF File (Adobe PDF File), 229KB - jmir_v16i7e171_app1.pdf]

Multimedia Appendix 2

Summary of included studies.

[PDF File (Adobe PDF File), 52KB - jmir_v16i7e171_app2.pdf]

Multimedia Appendix 3

Quality assessment of included studies.

[PDF File (Adobe PDF File), 33KB - jmir_v16i7e171_app3.pdf]

References

1. Boyd DM, Ellison NB. Social network sites: Definition, history, and scholarship. Journal of Computer‐Mediated Communication 2007 Oct;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
2. Smith T. Conference notes – The social media revolution. International Journal of Market Research 2009;51(4):559-561. [doi: 10.2501/S1470785309200773]
3. Eytan T, Benabio J, Golla V, Parikh R, Stein S. Social media and the health system. Perm J 2011;15(1):71-74 [FREE Full text] [Medline: 21505625]
4. Prasad B. Social media, health care, and social networking. Gastrointestinal endoscopy 2012 Oct 24;77(3):492-495 [FREE Full text] [doi: 10.1016/j.gie.2012.10.026]
5. Fielding NG, Lee RM, Hogan B. In: Fielding N, editor. The SAGE Handbook of Online Research Methods. London: Sage Publications Ltd; 2008.
6. Anderson-Butcher D, Lasseigne A, Ball A, Brzozowski M, Lehnert M, McCormick BL. Adolescent Weblog Use: Risky or Protective? Child Adolesc Soc Work J 2010 Jan 14;27(1):63-77. [doi: 10.1007/s10560-010-0193-x]
7. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368 [FREE Full text] [doi: 10.1377/hlthaff.28.2.361] [Medline: 19275991]
8. Silenzio VM, Duberstein PR, Tang W, Lu N, Tu X, Homan CM. Connecting the invisible dots: reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social networks. Soc Sci Med 2009 Aug;69(3):469-474 [FREE Full text] [doi: 10.1016/j.socscimed.2009.05.029] [Medline: 19540641]
9. Dissemination CFRA. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare. York: University of York NHS Centre for Reviews & Dissemination; 2009.

XSL•FO

RenderX

10.   Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Reprint--preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Phys Ther 2009 Sep;89(9):873-880 [FREE Full text] [Medline: 19723669]

11.   Levine D, Madsen A, Wright E, Barar RE, Santelli J, Bull S. Formative research on MySpace: online methods to engage hard-to-reach populations. J Health Commun 2011 Apr;16(4):448-454. [doi: 10.1080/10810730.2010.546486] [Medline: 21391040]

12.   Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012 Feb;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978]

13.   Heather J. Social media and health: Social support and social capital on pregnancy-related social networking sites, in Dissertation Abstracts International Section A: Humanities and Social Sciences. Los Angeles, California: University of Southern California; 2010.

14.   Woolley P, Peterson M. Efficacy of a Health-Related Facebook Social Network Site on Health-Seeking Behaviors. Social Marketing Quarterly 2012 May 02;18(1):29-39. [doi: 10.1177/1524500411435481]

15.   Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]

16.   Lohse B. Facebook is an effective strategy to recruit low-income women to online nutrition education. J Nutr Educ Behav 2013 Jan;45(1):69-76. [doi: 10.1016/j.jneb.2012.06.006] [Medline: 23305805]

17.   Lord S, Brevard J, Budman S. Connecting to young adults: an online social network survey of beliefs and attitudes associated with prescription opioid misuse among college students. Subst Use Misuse 2011;46(1):66-76. [doi: 10.3109/10826084.2011.521371] [Medline: 21190407]

18.   Shindel AW, Rowen TS, Lin TC, Li CS, Robertson PA, Breyer BN. An Internet survey of demographic and health factors associated with risk of sexual dysfunction in women who have sex with women. J Sex Med 2012 May;9(5):1261-1271 [FREE Full text] [doi: 10.1111/j.1743-6109.2012.02659.x] [Medline: 22375801]

19.   Cucchetti A, Zanello M, Bigonzi E, Pellegrini S, Cescon M, Ercolani G, et al. The use of social networking to explore knowledge and attitudes toward organ donation in Italy. Minerva Anestesiol 2012 Oct;78(10):1109-1116. [Medline: 23059515]

20.   Marsden PV, Wright JD. Handbook of survey research. In: Handbook of Survey Research, Second Edition. Bingley: Emerald Publishing Group Limited; 2010.

21.   Strasser SM, Smith M, Pendrick-Denney D, Boos-Beddington S, Chen K, McCarty F. Feasibility study of social media to reduce intimate partner violence among gay men in Metro Atlanta, Georgia. West J Emerg Med 2012 Aug;13(3):298-304 [FREE Full text] [doi: 10.5811/westjem.2012.3.11783]

Abbreviations

MSM: men who have sex with men
SNS: social network sites
WSW: women who have sex with women

XSL·FO

RenderX

Review

Assessing the Applicability of E-Therapies for Depression, Anxiety, and Other Mood Disorders Among Lesbians and Gay Men: Analysis of 24 Web- and Mobile Phone-Based Self-Help Interventions

Tomas Rozbroj[1], BA (Hons); Anthony Lyons[1], PhD; Marian Pitts[1], PhD; Anne Mitchell[1], MA; Helen Christensen[2], PhD

[1]Australian Research Centre in Sex, Health and Society, Faculty of Health Sciences, La Trobe University, Melbourne, Australia
[2]Black Dog Institute, Sydney, Australia

Corresponding Author:
Tomas Rozbroj, BA (Hons)
Australian Research Centre in Sex, Health and Society
Faculty of Health Sciences
La Trobe University
215
Franklin Street
Melbourne, 3000
Australia
Phone: 61 3 9479 8765
Fax: 61 3 9479 8711
Email: t.rozbroj@latrobe.edu.au

Abstract

Background: Lesbians and gay men have disproportionately high rates of depression and anxiety, and report lower satisfaction with treatments. In part, this may be because many health care options marginalize them by assuming heterosexuality, or misunderstand and fail to respond to the challenges specifically faced by these groups. E-therapies have particular potential to respond to the mental health needs of lesbians and gay men, but there is little research to determine whether they do so, or how they might be improved.

Objective: We sought to examine the applicability of existing mental health e-therapies for lesbians and gay men.

Methods: We reviewed 24 Web- and mobile phone-based e-therapies and assessed their performance in eight key areas, including the use of inclusive language and content and whether they addressed mental health stressors for lesbians and gay men, such as experiences of stigma related to their sexual orientation, coming out, and relationship issues that are specific to lesbians and gay men.

Results: We found that e-therapies seldom addressed these stressors. Furthermore, 58% (14/24) of therapies contained instances that assumed or suggested the user was heterosexual, with instances especially prevalent among better-evidenced programs.

Conclusions: Our findings, and a detailed review protocol presented in this article, may be used as guides for the future development of mental health e-therapies to better accommodate the needs of lesbians and gay men.

(J Med Internet Res 2014;16(7):e166)   doi:10.2196/jmir.3529

Introduction

Over the past two decades, Internet self-help therapy, hereafter referred to as "e-therapy", has rapidly become established as a recognized therapeutic approach in the treatment of depression and anxiety disorders [1-5]. E-therapy is also increasingly being delivered via mobile phone applications [6,7]. E-therapy's effectiveness as a general mode of therapy is now supported by a considerable body of research [3,8-11]. Individual e-therapies

XSL•FO
RenderX

within this health care mode have been tested to varying degrees. For example, MoodGym (Australia) has been examined in 17 research trials and seven randomized controlled trials [12], while some other currently available therapies have yet to be supported by research evidence.

E-therapies are an attractive option because they are relatively cost-effective [4,10,13-15], accessible, and able to maintain user anonymity [4,10,15]. In addition, e-therapies have been identified as particularly suitable for use by marginalized populations [13,16], such as rural persons [5] or same-sex attracted persons [5,16]. The use of e-therapies is now promoted by peak health care bodies, such as the National Health Service (NHS) in the United Kingdom [17,18], the Australian Psychological Society (APS) in Australia [3], and the Ministry of Health in New Zealand [19]. The NHS in particular has supported e-therapies by incorporating them into the United Kingdom's health care strategy and subsidizing their use [17,20]. The net outcome is that e-therapies are not only on the rise, but are now recognized as an integral constituent of future health care solutions of Australia, the United Kingdom, and New Zealand, with therapies also common in many other regions, including the United States and parts of Europe [19,21].

While e-therapies are demonstrably an important tool in addressing mental health issues, most are designed for the population in general. Little work has been done to evaluate whether their content and language accounts for and meets the needs of individuals who identify as lesbian or gay. In Australia, homosexual and bisexual persons are three times as likely to experience depression and twice as likely to experience anxiety as the general population [22]. They are less likely to seek treatment, and when they do, they tend to have considerable concerns about experiencing discrimination [23]. Previous research, such as that grounded in Minority Stress Theory, has shown that health care systems often assume heterosexuality, and that this can have adverse mental health effects for same-sex attracted persons [24-26]. These negative effects not only arise from overt discrimination, such as doctors treating patients differently, but may also arise from the incongruence experienced by interacting with social structures that do not take the minority group into account, where the group is invisible to its language and design [27-29]. Thus, an e-therapy that replicates this incongruence might inadvertently contribute to minority stress, perhaps resulting in some lesbians and gay men disconnecting with a treatment.

Furthermore, it is insufficient for health care solutions to simply acknowledge sexual diversity. Lesbians and gay men are known to experience a unique set of mental health challenges that require tailored resources [29-32], such as dealing with discrimination and other forms of stigma, "coming out" to family and friends, and managing the process of concealing and disclosing their sexual orientation at work, in social settings, and other facets of life [28-31]. While avoiding language and content that assume heterosexuality on the part of the user is important, content that specifically targets these and other challenges is needed to sufficiently cater to the mental health needs of lesbians and gay men [29,33-36].

At present, it is not known the extent to which existing e-therapies avoid an assumption of heterosexuality in their language and other content (such as imagery depicting solely heterosexual persons and survey questions that could only be applicable to heterosexual partnering), or the degree to which they specifically address the mental health challenges of lesbians and gay men. This article responds to some of these questions by presenting findings from a review we conducted of a large number of widely accessible English language e-therapies. The main aim of the review was to provide an analysis of the current field of e-therapies with regard to accommodating the needs of lesbians and gay men and, in doing so, to identify areas that might be considered for improvement in the future development of programs. While some findings presented in this paper may be useful to tailoring content to the needs of other sexual minorities, such as bisexual, transgender, or asexual persons, it was decided that each of these groups is sufficiently different to hold unique sets of needs with regard to health and experiences of discrimination [37,38], and each merit a study tailored to them. Given that such a scope was beyond this project, we have focused our enquiry explicitly on the needs of lesbians and gay men as a starting point.

Methods

Selection of E-Therapies

We included both Web-based interventions and app-based interventions. A Web-based intervention is defined as "a primarily self-guided intervention program that is executed by means of a prescriptive online program operated through a website and used by consumers seeking health and mental health related assistance" [39]. An app-based intervention is essentially the same as a Web-based intervention except that it is operated through a mobile phone application using phone memory and/or the Internet.

Given that the number of Web-based interventions and app-based interventions is large and rapidly growing, we focused on those that were provided in English and were most likely to be effective and widely accessible for the prevention or treatment of depression and anxiety. To build this sample, we used four selection criteria; an intervention was included if it met all four criteria. First, the intervention needed to provide a recognized therapeutic modality, as categorized by the Centre for Mental Health Research, which refers to therapies that engender mood/behavior change via the application of evidence-based methods, such as cognitive behavior therapy (CBT), acceptance and commitment therapy (ACT), and Narrative Therapy. This eliminated general information websites, diary apps, and other such content that would merit a separate enquiry and that did not match our aim of scrutinizing structured therapies. Second, the intervention had to target depression and/or anxiety disorders. This encompassed generalized depression, bipolar disorder, post-traumatic stress disorder, generalized anxiety disorder, social anxiety disorder, panic disorder, obsessive-compulsive disorder, and social phobia. Third, the intervention had to be open access and therefore more easily accessible than paid therapies. Therapies that are free to the residents of a particular country, such as

those subsidized by the NHS for citizens of the United Kingdom, were considered open access. Finally, the intervention needed to be in English. We did not have sufficient resources to translate non-English language e-therapies.

We used Beacon to source our sample. Beacon is a comprehensive database of e-health publications from around the world that is managed and regularly updated by the Centre for Mental Health Research at the Australian National University [12,40]. The database is compiled by a panel of experts in the field of mental e-health therapies and provides categorization and ranking of e-therapies based on expert evaluation from the panel. It stipulates classifications for all four of our selection criteria listed above: the disorder/s covered by each therapy, whether the therapy uses a recognized therapeutic modality, whether it is open access, and the language/s in which the therapy is delivered.

All Beacon-listed e-therapies were evaluated for selection on the basis of our selection criteria. The final sample was determined on 1st November, 2013. In all, 28 e-therapies met the selection criteria. However, we were unable to obtain researcher access to analyze four of the selected e-therapies. Thus, our final sample consisted of 24 e-therapies, which comprised 20 Web-based interventions and 4 app-based intervention. Following final selection, a second check was performed on the Beacon database to ensure that no errors were made in the selection process.

Procedure

Our measures can be divided into "attributes" and "content domains". Attributes are classifying characteristics of e-therapies, such as their length, structure type, and whether they have research evidence to support their effectiveness. Content domains are measures to assess how appropriate therapies are for lesbians and gay men.

The Beacon database provided much of the attribute information for the 24 therapies in our sample. We focused on the length of the therapy (short, medium, long; as categorized by Beacon), extent of research evidence that supported the therapy's effectiveness as measured by number of randomized controlled trials (RCTs) and other studies, country of origin of the therapy, and whether the therapy was a Web-based intervention or an app-based intervention. In addition, we also noted whether or not the e-therapy delivered content using scenarios: stories and examples involving characters. This last attribute was chosen given that e-therapies that contain scenarios have the added challenge of developing scenarios that are inclusive; for

example, providing stories that are not limited to heterosexual relationships.

We then analyzed each therapy according to eight content domains (see Table 1). These domains cover key areas in which e-therapies might feasibly be tailored to accommodate or appeal to lesbians and gay men. They were derived from previous research that has identified inequalities and challenges faced by lesbians and gay men [16,27-29,35,36,41,42], including issues of lesbians and gay men feeling invisible or left out in health care provisions (eg, as a result of the practitioner overtly assuming they are heterosexual), and inadequate provision of resources that are suited to the needs and experiences of lesbians and gay men, which include coming out (eg, disclosing sexual identity to family or colleagues), same-sex relationships, and lack of appropriate references to helplines. Furthermore, we responded to a range of stigma-related challenges as described by Minority Stress Theory, such as discrimination, prejudice, fear of discrimination, and internalized stigma [25]. In this case, internalized stigma refers to lesbians or gay men internalizing or adopting negative attitudes that others may have about lesbians or gay men. We refer to these stigma-related challenges under the broad term "homonegativity". We also analyzed whether therapies avoided instances that assumed or suggested the user was heterosexual. This involved assessing the appropriateness of language, such as whether a therapy referred to one's "spouse", which signifies marriage between a man and a woman. It also involved assessing the range of examples and scenarios used to deliver content within a therapy, and in particular whether all examples that indicated sexuality only depicted heterosexuality (eg, wives and husbands).

E-therapies were rated by the first author of this article on each content domain with a "yes" or "no", according to the presence or absence of content. Detailed notes were also taken to justify the rating and to assist with checking inter-rater reliability. Results were recorded in a spreadsheet and paired with analytical notes and references to screenshots and excerpts, and the date when the analysis for each e-therapy was conducted. The reliability of ratings was checked by a second independent researcher, who was not otherwise involved in the research and is not an author of this article. This independent researcher rated 20% of the sample, totaling 5 Web-based interventions. A 97.5% agreement was reached between the original ratings and those of the independent researcher. There was only one disagreement, which was resolved easily through discussion and followed up by further relevant checking to ensure the error was an isolated one. In all, inter-rater reliability scores indicated a high level of confidence in the accuracy of the original ratings.

Table 1. Eight content domains used to assess the applicability of 24 e-therapies for lesbians and gay men.

	Content domains
1	Referred to lesbians and gay men in the introductory section of the therapy
2	Explicitly addressed homonegativity
3	Explicitly addressed coming out
4	Explicitly referred to same-sex relationships
5	Used imagery that depicted lesbians and/or gay men
6	Avoided instances that assumed or suggested the user was heterosexual
7	Provided references to mental health resources aimed at lesbians and gay men
8	Explicitly referred to lesbians and gay men in other ways not captured by the above

Results

Profile of E-Therapies

Table 2 summarizes attributes of the 24 e-therapies. A majority (67%, 16/24) were targeted to the treatment of depression and/or generalized anxiety disorder. Three-quarters (75%, 18/24) delivered content using scenarios. Just over half (54%, 13/24) were developed in Australia. This was in large part because fewer of those from other countries, particularly the United Kingdom and United States, were available for free and therefore these did not meet our criterion of open access. Half (50%, 12/24) were supported by research evidence and three of these were verified by randomized controlled trials. Almost all (92%, 22/24) therapies were categorized by Beacon as long.

Table 2. Profile of e-therapies (N=24).

E-therapy profile		n (%)
Therapy type		
	Web-based intervention	20 (83)
	App-based intervention	4 (17)
Disorder type [a]		
	Depression	12 (50)
	Bipolar	0 (0)
	Generalized Anxiety Disorder	11 (46)
	Social Anxiety Disorder	3 (13)
	Panic Disorder	3 (13)
	Post-Traumatic Stress Disorder	4 (17)
	Phobia	2 (8)
	Obsessive Compulsive Disorder	1 (4)
Content delivery method		
	Scenario	18 (75)
	Non-scenario	6 (25)
Origin		
	Australia	13 (54)
	United States	5 (21)
	United Kingdom	3 (13)
	New Zealand	1 (4)
	United States/Canada[b]	1 (4)
	Israel	1 (4)
Evidence rating (x/5) [c]		
	Two or more	3 (13)
	One	9 (38)
	Zero	12 (50)
Length [d]		
	Long	22 (92)
	Moderate	2 (8)
	Short	0 (0)

[a]Therapies may cater to multiple disorders, therefore total exceeds N=24 (100%).

[b]Joint collaboration between United States and Canada.

[c]Evidence rating is the score awarded by Beacon to indicate the degree to which an e-therapy is supported by research evidence. Zero indicates no evidence or no evidence of effectiveness. One indicates some evidence but no evidence from randomized controlled trials. Two or higher indicates evidence of effectiveness, including from randomized controlled trials.

[d]Long: 5+ modules; Moderate: 3-5 modules; Short: 1-2 modules.

Overall Ratings

Table 3 displays numbers and percentages of therapies that scored positively in the eight content domains that we examined in our review. In all, few therapies scored in content domains 1-5. These domains concerned topics of coming out, homonegativity, same-sex relationships, visually depicting lesbians and/or gay men, and acknowledging lesbians and gay men in the introductory sections of a therapy. Examples of this content included forum articles discussing experiences of coming out to parents, dealing with sexuality-based abuse, and experiences of entering the same-sex dating scene. Four (17%, 4/24) therapies were found to have references to mental health resources aimed at lesbians and gay men. In all cases, these comprised a list of crisis referral services aimed at lesbians, gay men, and other sexual minorities. Ten therapies (42%, 10/24)

were found to avoid instances that assumed or suggested users were heterosexual. Instances where e-therapies that failed to do this included referring to a partner as a "spouse" (a term that typically denotes heterosexual partnering) and imagery that depicted only heterosexual relationships and/or nuclear families.

One e-therapy, Big White Wall (BWW), accounted for the majority of domain scores. BWW scored in domains 1-5 and 8, while none of the other therapies scored in more than two content domains. BWW is comprised of both professionally-created and user-created content. The latter was largely responsible for the positive scores. Professional content actually failed to avoid instances that assumed user heterosexuality and only scored positively in one other content

domain, when it mentioned same-sex attracted persons twice: in an interview transcript about "managing your differences" and in a review of a book that included a gay character. The count of user-generated same-sex orientation specific content was comparably enormous. When we entered keywords into BWW's search box on the 20th November, we found 120 artworks and 129 forum results for the word "gay" and 56 and 57 respectively for the word "lesbian". Other keywords such as "queer" and "homosexual" returned a smaller number of results. The primarily inclusive nature of these posts were confirmed on closer scrutiny, which revealed content about coming out, family issues relating to sexuality, inclusivity, resilience, artworks with rainbow flags, gay couples, and even a picture protesting against Uganda's recent anti-homosexuality Act.

Table 3. Numbers and percentages of e-therapies that scored in each content domain (N=24).

	Content domains	Yes, n (%)
1	Referred to lesbians and gay men in the introductory section of the therapy	0 (0)
2	Explicitly addressed homonegativity	1 (4)
3	Explicitly addressed coming out	2 (8)
4	Explicitly referred to same-sex relationships	1 (4)
5	Used imagery that depicted lesbians and/or gay men	2 (8)
6	Avoided instances that assumed or suggested the user was heterosexual	10 (42)
7	Provided references to mental health resources aimed at lesbians and gay men	4 (17)
8	Explicitly referred to lesbians and gay men in other ways not captured by the above	1 (4)

Ratings According to Content Delivery Method and Evidence Rating

Given that e-therapies that use scenarios (eg, stories involving characters) have the added challenge of developing scenarios that are inclusive of marginalized populations, and may be more appealing to users, we examined whether e-therapies that scored in each of the eight domains were among those that used scenarios. All of the therapies that had content in domains 1-5 and 7-8 were found to use scenarios. Thus, none of the non-scenario therapies scored in these domains. In contrast, only four of the 10 therapies that scored in content domain 6 (ie, avoided instances that assumed user heterosexuality) used scenarios while the remaining six therapies did not use scenarios. Thus, all non-scenario therapies scored in this domain.

Assuming that better-evidenced e-therapies are also more likely to appeal to users, we further examined whether therapies that scored in each of the eight content domains were among those with research evidence. In all, therapies that scored in the content domains were mostly those supported by research evidence. In particular, all of the therapies that had content in domains 1-5 and 7-8 had an evidence rating of 1 or 2, which means that they were supported by at least some research evidence. Of the 10 therapies that scored in content domain 6 (ie, avoided instances that assumed user heterosexuality), five (50%) had an evidence rating of 1 or 2 and the remaining five (50%) had an evidence rating of zero. Thus, e-therapies with an evidence rating of zero did not score in any of the content domains except for content domain 6. The three e-therapies with an evidence rating of 2 or higher (ie, were supported by

evidence from randomized controlled trials) each scored in one domain, with one scoring in domain 1, one in domain 5, and one in domain 6.

As suggested by the above results, there was considerable overlap between therapies that were supported by research evidence and those that used scenarios. Specifically, of the 18 e-therapies that used scenarios, three (17%) had evidence ratings of 2 or higher, eight a rating of 1 (44%), and seven (39%) a rating of zero. Of the six e-therapies that did not use scenarios, none had evidence ratings of 2 or higher, only one (17%) had an evidence rating of 1, and five (83%) had a rating of zero.

Discussion

Principal Findings

Overall, e-therapies seldom catered to the needs of lesbians and gay men. Most did not include content that explicitly covered key experiences such as coming out or being in a same-sex relationship. For the most part, the language used did not account for same-sex attracted clients. Further, more than half the e-therapies used content that assumed or suggested user heterosexuality. It would appear that the experiences of lesbians and gay men were seldom considered in the development of these e-therapies. This corroborates with past research in this area, which has found that no existing computerized cognitive behavioral therapy (cCBT) programs address challenges specific to lesbians and gay men [16]. It also reinforces broader work that demonstrates a shortage of mental health care that caters to the needs of same-sex attracted populations [16,28,31]. This is despite a comprehensive and growing body of research that

XSL•FO
RenderX

shows that therapy modes that fail to do so can alienate lesbians and gay men and lead to diminished therapeutic outcomes [27,29-31,34].

Interestingly, whether or not a therapy used scenarios (ie, stories and examples involving other characters) resulted in markedly different content domain scores. All of the non-scenario therapies did not assume sexual orientation in their language and content, but they also did not record a positive score across any of the other content domains. On the other hand, only 42% of scenario therapies avoided instances that assumed user heterosexuality, but some scored positively in other content domains. This is attributable to distinctions in the structural characteristics of the two therapy types. All therapies—both scenario and non-scenario—were careful to address their audience in neutral language, to cater to both sexes if nothing else. The locus of content that assumed heterosexuality was in the extra layer of content in scenario therapies. These scenarios included stories about nuclear families, heterosexual dating, and so forth. Audio-visual content often exemplified the text and was also problematic. This content tended to be prominent and pervasive in therapies and failed to present non-heterosexual alternatives. Non-scenario therapies avoided the issue of assuming heterosexuality by not having this layer. They also tended to present content in a broader way; where a scenario therapy may discuss the financial strain on John's marriage, a non-scenario therapy may discuss "building positive relationships". Issues such as coming out or homonegativity were too narrow to fall within the scope of the latter, just as content specifically about heterosexual relationships was too narrow to be explicitly addressed. This contrasted with scenario therapies, which frequently covered heterosexual issues and addressed specific problems such as dating or partnering. Thus, on the one hand, non-scenario therapies expressed content broadly, which had the effect of not excluding specific sexualities but also not dealing with sexuality-specific content. On the other hand, scenario therapies dealt with sexuality-specific content but largely omitted same-sex attraction.

In our sample, the scenario therapies were more common and scored better on Beacon ratings of research evidence than non-scenario therapies. In fact, the majority were supported by research evidence, as compared to only one-sixth of non-scenario therapies. Scenario therapies appear to be the more prominent therapy type, certainly within our sample, and given that individuals are more likely to be referred to evidence-based e-therapies by health professionals, the low content domain scores and high frequency of heteronormative language in scenario therapies may need attention in future development of programs.

As our results show, only one to two therapies included content that specifically addressed mental health challenges experienced by lesbians and gay men. Big White Wall (BWW) is the program that accounted for the majority of these cases. BWW is a UK-based therapy designed to help treat stress, depression, and general anxiety. Beacon described it as a service that "utilizes the principles of social networking" [12]. This is because while BWW offers a range of professionally created self-help material and guided activities, it also integrates

user-created social support groups composed of BWW clients. They can start forum discussions and express themselves through art and writing. The convergence of professional and user-generated content is integral in BWW, which is composed of polygenic content organized in a non-linear structure. This differentiates it from major therapies like MoodGym, Beating the Blues, and AnxietyOnline. Such a format may be significant for lesbians and gay men, as positive networks have been identified as core resilience drivers for same-sex attracted persons [43-45].

BWW's positive scores, and the fact that user content was largely responsible, highlights two important points. First, the level of user posts pertaining to same-sex attraction demonstrates the merit of incorporating social network-styled structures into e-therapies to assist in catering to lesbians and gay men. For the many therapies that cater to the general population, this may be an efficient way to provide at least some content that targets lesbians and gay men. However, the second point is that the stark contrast between the level of professional and user-generated content illustrates a disparity between supply and demand for content that is specific to lesbians and gay men. The volume of topics raised, many of which paralleled our content domains, demonstrate a need by lesbians and gay men for information and guidance that is not currently being delivered by professional content. It also highlights that such issues as coming out, family acceptance, and inclusivity are important parts of the mental health experiences of lesbians and gay men, and there is a demand for therapies to address them. One of the particular advantages of e-therapies as a delivery mode is their capacity to deliver tailored, targeted content, perhaps using adaptive logic. Harnessing this capacity by targeting specific content to a user according to their sexual orientation is one way of catering to lesbians and gay men within e-therapies, but so far this appears to have been underutilized.

Limitations and Future Directions

Findings of this review are limited to e-therapies we selected. While we chose selection criteria that enabled a diverse sample of e-therapies, they were nonetheless limited to those that used therapeutic modalities (eg, CBT, ACT), targeted depression and anxiety, were open access, and were in English. Language in particular is a notable limitation, as there are many instances of e-therapies in other languages. That said, we are confident that our sample was large and diverse enough for our findings to be broadly applicable to e-therapies, at least in English-speaking countries.

Our findings were also limited to e-therapies that were available on November 2013. The e-therapy landscape is changing rapidly, with existing e-therapies constantly being modified and new therapies being developed. For example, some e-therapies that would have fit our selection criteria were not included in the analysis because they were in research trial phases at the time. It should be noted that SPARX—an adventure computer game that treats depression—has recently been trialed in a "Rainbow" version for adolescents who are same-sex attracted or questioning [16,46]. This therapy was not available at the time of our analysis, and is only available on CD-ROM and not currently available online. App-based interventions are

XSL·FO

RenderX

propagating even more quickly, and we suspect that at the time of publication of this article there will be a significantly larger number of app-based interventions that might have fit our selection criteria. Finally, our findings and review protocol are limited to lesbians and gay men. We felt this limitation was necessary because issues faced by other sexual minorities, such as people who identify as bisexual or pansexual, are often different compared to lesbians and gay men [23,37,38] and may therefore need to be handled somewhat differently in e-therapies. In future, studies are needed to examine ways in which e-therapies might be further improved to accommodate the specific needs of other sexual minorities.

Despite its limitations, our work drew a relevant sketch of the deficiencies of current e-therapies in catering to lesbians and gay men, and may help to guide future projects to improve this. On the whole, it is clear that greater attention to the needs of lesbians and gay men is desirable in the design of e-therapies. Consideration should be given to the construction of language—verbal and audio-visual—that does not preclude the experiences of same-sex attracted persons. We acknowledge that developers of e-therapies that are intended for a general audience may be concerned that adding content specifically tailored to lesbians and gay men might be off-putting to predominantly heterosexual users. Tailoring to all eight of our content domains may not be appropriate for all therapies, but it may be valuable to consider whether minor changes such as including gay or lesbian-specific helplines or some acknowledgement of same-sex attracted persons may make the e-therapy more attractive for these populations. As noted earlier, using adaptive logic to tailor content to a user's sexual identity could be a particularly powerful way of ensuring content meets the needs of specific sexual identity groups without deterring other groups. In fact, our review protocol could be useful to developers as a tool to assess the inclusivity of their therapeutic programs, by repeating the steps outlined in our methodology to conduct a review of their own e-therapy and to identify potential areas where content could be improved or tailored. In terms of future research, we believe that further inquiry is needed from the perspectives of same-sex attracted persons to identify specific kinds of content that would make them feel more included and that address the specific mental health needs of these populations.

Conclusions

In this article, we presented the findings of a review of Web-based and app-based interventions for the prevention and treatment of depression and anxiety with regard to the degree to which they catered to lesbians and gay men. The majority did not address many of the key additional factors for depression and anxiety experienced by lesbians and gay men. They largely did not acknowledge lesbians and gay men, or address core issues like coming out, dealing with discrimination and prejudice, or same-sex relationships. Many of the therapies that used scenarios to deliver content, which tended to be the more prominent type of e-therapy in terms of numbers and evidence of effectiveness, also contained instances of language and content that assumed user heterosexuality. As we have outlined, past research indicates that many therapies may exclude lesbians and gay men and in doing so may inadvertently contribute to minority stress. This is particularly an issue given that these populations already experience comparatively high rates of mental health problems, are less likely to engage with therapy, and are known to express concerns around fears of discrimination. Our findings suggest that e-therapies could do more to address the needs of lesbians and gay men, thus enabling these populations to benefit more greatly from the rise of e-therapy for depression and anxiety and to therefore contribute to reducing disparities in mental health between heterosexual and non-heterosexual populations. To this end, we suggest that the review protocol developed for this study could be utilized by developers to help monitor and improve the applicability of e-therapies for lesbians and gay men.

Conflicts of Interest

None declared.

References

1. Cowpertwait L, Clarke D. Effectiveness of web-based psychological interventions for depression: A meta-analysis. Int J Ment Health Addiction 2013 Jan 18;11(2):247-268. [doi: 10.1007/s11469-012-9416-z]
2. Van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008 Mar;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]
3. Fuller T, Stokes D, Mathews R. Internet supported psychological interventions: A guide to navigating the online world of psychological programs. Australia: APS Professional Practice; 2013.
4. Foroushani PS, Schneider J, Assareh N. Meta-review of the effectiveness of computerised CBT in treating depression. BMC Psychiatry 2011;11:131 [FREE Full text] [doi: 10.1186/1471-244X-11-131] [Medline: 21838902]

XSL•FO

RenderX

5. Rochlen AB, Zack JS, Speyer C. Online therapy: review of relevant definitions, debates, and current empirical support. J Clin Psychol 2004 Mar;60(3):269-283. [doi: 10.1002/jclp.10263] [Medline: 14981791]

6. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, 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]

7. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]

8. Heilman RM, Kállay É, Miclea M. The role of computer-based psychotherapy in the treatment of anxiety disorders. Cognition, Brain, Behavior 2010;14(3):209-230.

9. Postel MG, de Haan HA, De Jong CA. E-therapy for mental health problems: a systematic review. Telemed J E Health 2008 Sep;14(7):707-714. [doi: 10.1089/tmj.2007.0111] [Medline: 18817501]

10. Abbott JM, Klein B, Ciechomski L. Best practices in online therapy. Journal of Technology in Human Services 2008 Jul 03;26(2-4):360-375. [doi: 10.1080/15228830802097257]

11. Christensen H. Computerised therapy for psychiatric disorders. Lancet 2007 Jul 14;370(9582):112-113. [doi: 10.1016/S0140-6736(07)61065-4] [Medline: 17630022]

12. Centre for Mental Health Research. 2014. Beacon URL: https://beacon.anu.edu.au/ [accessed 2014-06-30] [WebCite Cache ID 6Qi5ABInK]

13. Spurgeon JA, Wright JH. Computer-assisted cognitive-behavioral therapy. Curr Psychiatry Rep 2010 Dec;12(6):547-552. [doi: 10.1007/s11920-010-0152-4] [Medline: 20872100]

14. Christensen H, Griffiths K, Groves C, Korten A. Free range users and one hit wonders: community users of an Internet-based cognitive behaviour therapy program. Aust N Z J Psychiatry 2006 Jan;40(1):59-62. [doi: 10.1111/j.1440-1614.2006.01743.x] [Medline: 16403040]

15. Ybarra ML, Eaton WW. Internet-based mental health interventions. Ment Health Serv Res 2005 Jun;7(2):75-87. [doi: 10.1007/s11020-005-3779-8]

16. Lucassen MFG, Hatcher S, Stasiak K, Fleming T, Shepherd M, Merry SN. The views of lesbian, gay and bisexual youth regarding computerised self-help for depression: An exploratory study. Advances in Mental Health 2013 Oct;12(1):22-33. [doi: 10.5172/jamh.2013.12.1.22]

17. National Collaborating Centre for Mental Health. Depression: The NICE Guideline on the Treatment and Management of Depression in Adults (Updated Edition). Leicester, UK: RCPsych Publications; 2013.

18. Cavanagh K, Seccombe N, Lidbetter N, Bunnell D. Supported, service-user led, computerised cognitive behavioural therapy (CCBT) self-help clinics. Journal of Public Mental Health 2011;10(4):225-233. [doi: 10.1108/17465721111188241]

19. New Zealand Ministry of Health. SPARX e-therapy for young people launched URL: http://www.health.govt.nz/news-media/news-items/sparx-e-therapy-young-people-launched [accessed 2014-06-10] [WebCite Cache ID 6QDZ76ByT]

20. National Collaborating Centre for Primary Care. National Institute for Clinical Excellence. UK: University of Sheffield; 2007. Clinical guidelines for the management of anxiety: Management of anxiety (panic disorder, with or without agoraphobia and generalised anxiety disorder) in adults in primary, secondary and community care URL: http://publications.nice.org.uk/generalised-anxiety-disorder-and-panic-disorder-with-or-without-agoraphobia-in-adults-cg113/guidance [accessed 2014-06-30] [WebCite Cache ID 6QjLvugsT]

21. 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]

22. Australian Bureau of Statistics. National survey of mental health and wellbeing: Summary of results. Canberra: Australian Bureau of Statistics; 2007.

23. Leonard W, Pitts M, Mitchell A, Lyons A, Smith A, Patel S, et al. Private Lives 2: The second national survey of the health and wellbeing of gay, lesbian, bisexual and transgender (GLBT) Australians. Melbourne: Australian Research Centre in Sex, Health and Society; 2012.

24. Balsam KF, Beadnell B, Molina Y. The Daily Heterosexist Experiences Questionnaire: Measuring Minority Stress Among Lesbian, Gay, Bisexual, and Transgender Adults. Meas Eval Couns Dev 2013 Jan;46(1):3-25 [FREE Full text] [doi: 10.1177/0748175612449743] [Medline: 24058262]

25. Meyer IH. Prejudice and pride: Minority stress and mental health in gay men. New York: Columbia University; 1993.

26. Meyer IH. Minority stress and mental health in gay men. J Health Soc Behav 1995 Mar;36(1):38-56. [Medline: 7738327]

27. Meyer IH. Why lesbian, gay, bisexual, and transgender public health? Am J Public Health 2001 Jun;91(6):856-859. [doi: 10.2105/AJPH.91.6.856]

28. Saulnier CF. Deciding who to see: Lesbians discuss their preferences in health and mental health care providers. Social Work 2002 Oct 01;47(4):355-365. [doi: 10.1093/sw/47.4.355]

29. Murphy BC. Educating mental health professionals about gay and lesbian issues. J Homosex 1992 Apr;22(3-4):229-246. [doi: 10.1300/J082v22n03_10] [Medline: 1573259]

30. Peterson KJ. Health care for lesbians and gay men: confronting homophobia and heterosexism. New York: Haworth Press; 1996.

31. Morrison S, Dinkel S. Heterosexism and health care: a concept analysis. Nurs Forum 2012;47(2):123-130. [doi: 10.1111/j.1744-6198.2011.00243.x] [Medline: 22512770]

32. Lyons A, Pitts M, Grierson J. Growing old as a gay man: Psychosocial well-being of a sexual minority. Research on Aging 2012 May 03;35(3):275-295. [doi: 10.1177/0164027512445055]

33. Rostosky SS, Riggle EDB, Gray BE, Hatton RL. Minority stress experiences in committed same-sex couple relationships. Professional Psychology: Research and Practice 2007;38(4):392-400. [doi: 10.1037/0735-7028.38.4.392]

34. Röndahl G. Lesbians' and gay men's narratives about attitudes in nursing. Scand J Caring Sci 2009 Mar;23(1):146-152. [doi: 10.1111/j.1471-6712.2008.00603.x] [Medline: 19192241]

35. Berg MB, Mimiaga MJ, Safren SA. Mental health concerns of gay and bisexual men seeking mental health services. J Homosex 2008 May;54(3):293-306. [doi: 10.1080/00918360801982215] [Medline: 18825866]

36. Hoffman ND, Freeman K, Swann S. Healthcare preferences of lesbian, gay, bisexual, transgender and questioning youth. J Adolesc Health 2009 Sep;45(3):222-229 [FREE Full text] [doi: 10.1016/j.jadohealth.2009.01.009] [Medline: 19699417]

37. Friedman MR, Stall R, Silvestre AJ, Mustanski B, Shoptaw S, Surkan PJ, et al. Stuck in the middle: longitudinal HIV-related health disparities among men who have sex with men and women. J Acquir Immune Defic Syndr 2014 Jun 1;66(2):213-220. [doi: 10.1097/QAI.0000000000000143] [Medline: 24662298]

38. Kuyper L, Fokkema T. Minority stress and mental health among Dutch LGBs: examination of differences between sex and sexual orientation. J Couns Psychol 2011 Apr;58(2):222-233. [doi: 10.1037/a0022688] [Medline: 21401219]

39. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]

40. Christensen H, Murray K, Calear AL, Bennett K, Bennett A, Griffiths KM. Beacon: a web portal to high-quality mental health websites for use by health professionals and the public. Med J Aust 2010 Jun 7;192(11 Suppl):S40-S44. [Medline: 20528708]

41. Grabski B, Iniewicz G, Mijas M. Mental health of homosexual and bisexual people--review of research and presentation of phenomenon. Psychiatr Pol 2012;46(4):637-647. [Medline: 23214165]

42. Radkowsky M, Siegel LJ. The gay adolescent: Stressors, adaptations, and psychosocial interventions. Clinical Psychology Review 1997 Jan;17(2):191-216. [doi: 10.1016/S0272-7358(97)00007-X]

43. Lyons A, Pitts M, Grierson J. Sense of coherence as a protective factor for psychological distress among gay men: a prospective cohort study. Anxiety Stress Coping 2014 Feb 27:1-16. [doi: 10.1080/10615806.2014.887071] [Medline: 24456279]

44. Lyons A, Pitts M, Grierson J. Factors related to positive mental health in a stigmatized minority: an investigation of older gay men. J Aging Health 2013 Oct;25(7):1159-1181. [doi: 10.1177/0898264313495562] [Medline: 23860179]

45. Szymanski DM. Examining potential moderators of the link between heterosexist events and gay and bisexual men's psychological distress. Journal of Counseling Psychology 2009;56(1):142-151. [doi: 10.1037/0022-0167.56.1.142]

46. Lucassen MF, Merry SN, Hatcher S, Frampton CM. Rainbow SPARX: A novel approach to addressing depression in sexual minority youth. Cognitive and Behavioral Practice 2014 Apr:1. [doi: 10.1016/j.cbpra.2013.12.008]

Abbreviations

ACT: acceptance and commitment therapy
APS: Australian Psychological Society
BWW: Big White Wall e-therapy
CBT: cognitive behavior therapy
CCBT: computerized cognitive behavioral therapy
NHS: National Health Service
RCT: randomized controlled trial

Review

Employing Computers for the Recruitment into Clinical Trials: A Comprehensive Systematic Review

Felix Köpcke[1], BASc, Dipl-Ing; Hans-Ulrich Prokosch[2], Dr rer biol hum, PhD

[1]Center for Information and Communication, University Hospital Erlangen, Erlangen, Germany

[2]Department for Medical Informatics, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

Corresponding Author:

Hans-Ulrich Prokosch, Dr rer biol hum, PhD
Department for Medical Informatics
Friedrich-Alexander University Erlangen-Nuremberg
Lehrstuhl für Medizinische Informatik
Wetterkreuz 13
Erlangen, 91058
Germany
Phone: 49 91318526720
Fax: 49 91318526754
Email: Hans-Ulrich.Prokosch@uk-erlangen.de

Abstract

Background: Medical progress depends on the evaluation of new diagnostic and therapeutic interventions within clinical trials. Clinical trial recruitment support systems (CTRSS) aim to improve the recruitment process in terms of effectiveness and efficiency.

Objective: The goals were to (1) create an overview of all CTRSS reported until the end of 2013, (2) find and describe similarities in design, (3) theorize on the reasons for different approaches, and (4) examine whether projects were able to illustrate the impact of CTRSS.

Methods: We searched PubMed titles, abstracts, and keywords for terms related to CTRSS research. Query results were classified according to clinical context, workflow integration, knowledge and data sources, reasoning algorithm, and outcome.

Results: A total of 101 papers on 79 different systems were found. Most lacked details in one or more categories. There were 3 different CTRSS that dominated: (1) systems for the retrospective identification of trial participants based on existing clinical data, typically through Structured Query Language (SQL) queries on relational databases, (2) systems that monitored the appearance of a key event of an existing health information technology component in which the occurrence of the event caused a comprehensive eligibility test for a patient or was directly communicated to the researcher, and (3) independent systems that required a user to enter patient data into an interface to trigger an eligibility assessment. Although the treating physician was required to act for the patient in older systems, it is now becoming increasingly popular to offer this possibility directly to the patient.

Conclusions: Many CTRSS are designed to fit the existing infrastructure of a clinical care provider or the particularities of a trial. We conclude that the success of a CTRSS depends more on its successful workflow integration than on sophisticated reasoning and data processing algorithms. Furthermore, some of the most recent literature suggest that an increase in recruited patients and improvements in recruitment efficiency can be expected, although the former will depend on the error rate of the recruitment process being replaced. Finally, to increase the quality of future CTRSS reports, we propose a checklist of items that should be included.

Introduction

Medical progress depends on the evaluation of new diagnostic and therapeutic interventions within clinical trials. The value of each clinical trial depends on the successful recruitment of patients within a limited time frame. The number of participants must be sufficiently large to allow for scientifically and statistically valid analysis. Unfortunately, many trials experience

XSL•FO
RenderX

gaps between initially planned and finally achieved participant numbers or they need to prolong their recruitment period. Slow recruitment delays medical progress and leads to unnecessarily high study costs [1-3].

The main stakeholders in the recruitment process are the patient, the treating physician, the study nurse, and the principal investigator. But when it comes to the details of how responsibilities and tasks are distributed and how stakeholders interact with one another, recruitment processes start to show large variability. These specifics are influenced by a multitude of factors, including whether the trial is prospective or retrospective, the number of patients to be screened, the fraction of potential participants among the screened patients, the number of participating clinics, the urgency of recruiting a patient after discovering eligibility, the local data protection laws, the available funds or the organization, and infrastructure of the clinical institutions which pursue the trial.

Because of this variability in the recruitment processes, numerous reasons for failure to include sufficient participants into a trial were found [4-6]. On the most abstract level, these are overoptimistic feasibility estimations of future eligible patient numbers [7,8], the inability to motivate physicians to approach their patients [9-12], and the inability to motivate patients to participate [13,14].

Following increased levels of patient data capture in digital systems and the advent of clinical decision support systems, the early 1990s also saw the use of computers for matching patients and trial protocols. These clinical trial recruitment support systems (CTRSS) aim to solve the issue of false feasibility estimations, to generate a positive impact on the treating physicians' enrollment efforts, and to reduce the resources required to set up a successful recruitment process. Although many CTRSS have been proposed, the problems in recruitment persist [15,16].

In this context, Cuggia et al [17] raised the question "What significant work has been carried out toward automating patient recruitment?" and reviewed the literature published between 1998 and October 2009. They found a comparatively small number of papers related to 28 distinct CTRSS. Most of these projects had focused on the technological feasibility of the search algorithm and neglected assessments of the system's impact on recruitment in real-life scenarios. Cuggia et al concluded "that the automatic recruitment issue is still open" and that in 2009 it was still "difficult to make any strong statements about how effective automatic recruitment is, or about what makes a good decision support system for clinical trial recruitment."

Since then, CTRSS have become even more popular. Many independent institutions have tackled the challenge to improve their local recruitment processes. Large European collaborations, such as Electronic Health Records for Clinical Research (EHR4CR) [18], and national collaborations, for example in Germany [19], have been initiated to create information technology (IT)-supported patient recruitment architectures and platforms. For the related but broader challenge of extracting meaningful patient information from electronic health record (EHR) data, a plethora of publications have been published in recent years and the term *patient phenotyping* has been coined [20]. Recently, Shivade et al [21] presented a review on phenotyping techniques. They observed "a rise in the number of studies associated with cohort identification using electronic medical records."

The rapidly growing knowledge about and the importance of electronic patient recruitment systems warrants a new review of the existing literature. Our objectives were to (1) create an overview of all papers published until the end of 2013, (2) find and describe similarities in CTRSS design, (3) discuss the reasons for different approaches, and (4) examine whether new projects were able to illustrate the impact of CTRSS.

Methods

Search Strategy

One of the authors (FK) searched the database PubMed with 2 queries. The first query contained keywords for publication titles and Medical Subject Headings (MeSH) terms. Because most recent articles were not yet completely indexed with MeSH terms, a second query performed a more profound keyword search in all fields. Neither query was limited to a specific time period:

1. PubMed query 1: ("clinical trial"[Title] OR "clinical trials"[Title] OR "Clinical Trials as Topic"[MESH]) AND ("eligibility"[Title] OR "identification"[Title] OR "recruitment"[Title] OR "Patient Selection"[MESH] OR "cohort"[Title] OR "accrual"[Title] OR "enrollment"[Title] OR "enrolment" [Title] OR "screening"[Title]) AND ("electronic"[Title] OR "computer"[Title] OR "software"[Title] OR "Decision Making, Computer-Assisted"[MESH] OR "Decision Support Systems, Clinical"[Mesh] OR "Medical Records Systems, Computerized"[Mesh])

2. PubMed query 2: ("clinical trial"[All Fields] OR "clinical trials"[All Fields]) AND ("eligibility"[All Fields] OR "identification"[All Fields] OR "recruitment"[All Fields] OR "accrual"[All Fields] OR "enrollment"[All Fields] OR "enrolment"[All Fields] OR "screening"[All fields]) AND ("participants"[All Fields] OR "cohort"[All fields] OR "patients"[All Fields]) AND ("electronic"[All fields] OR "computer"[All fields] OR "software"[All fields] OR "automatic"[All Fields])

Both queries were executed on January 15, 2014. After removing all duplicates from the combined result sets of both queries, FK screened titles and abstracts for the inclusion criteria. We then tried to obtain the full text of all included articles for a second screening. Finally, FK reviewed all references of the included manuscripts for additional articles. In case of uncertainty about the inclusion of an article, it was discussed with HUP for a final decision.

Inclusion Criteria

Our review covers primary research articles and conference proceedings on computer systems that compared patient data and eligibility criteria of a clinical trial to identify either potential participants for a given trial or suitable trials for a given patient. The system must have employed a computer to

determine patient eligibility; that is, the utilization of electronically captured data was insufficient if the matchmaking process itself was done manually (eg, [22,23]). Manual processes before and after eligibility determination were otherwise accepted. Articles on the construction and processing of eligibility criteria, although closely tied to the construction and usage of CTRSS, were not part of this review (eg, [24,25]). Although technically the same, we also excluded decision support systems that identified patients for other purposes than clinical trials recruitment (eg, for diagnosing [26] or phenotyping [27]).

Classification

The classification of CTRSS was roughly based on that of a previous review by Cuggia et al [17] to render results comparable with one another. They included (1) the clinical context or setting to which the system was deployed, (2) the manner of integration into the existing clinical or recruitment workflow, (3) the source and format of patient data and eligibility criteria, (4) the reasoning method employed to derive eligible patients, and (5) the outcome obtained by the system's application to one or more clinical trials.

Results

Included Studies

The 2 PubMed queries together yielded 1693 articles. A total of 1581 articles were removed from the literature pool based on their titles and abstracts. After removal of 8 articles that could not be obtained as full text and 21 duplicates, we arrived at 83 distinct articles, of which 60 were included in the qualitative analysis after review. In all, 5 of the excluded articles described other supportive measures for trial recruitment, 4 were deemed nonscientific (eg, commentaries), 6 described manual systems or the mode of eligibility determination was not clearly stated, 3 constituted general contributions without a relation to a specific CTRSS, 3 focused on the representation of eligibility criteria in a computable format, and 2 articles dealt with other topics (eg, phenotyping, personalized medicine). We obtained 41 additional articles through references and arrived at a final pool of 101 articles [3,28-127] on 79 different systems. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram [128] in Figure 1 shows the different phases of the article selection process.

**Figure 1.** Flow diagram on the process of literature selection.
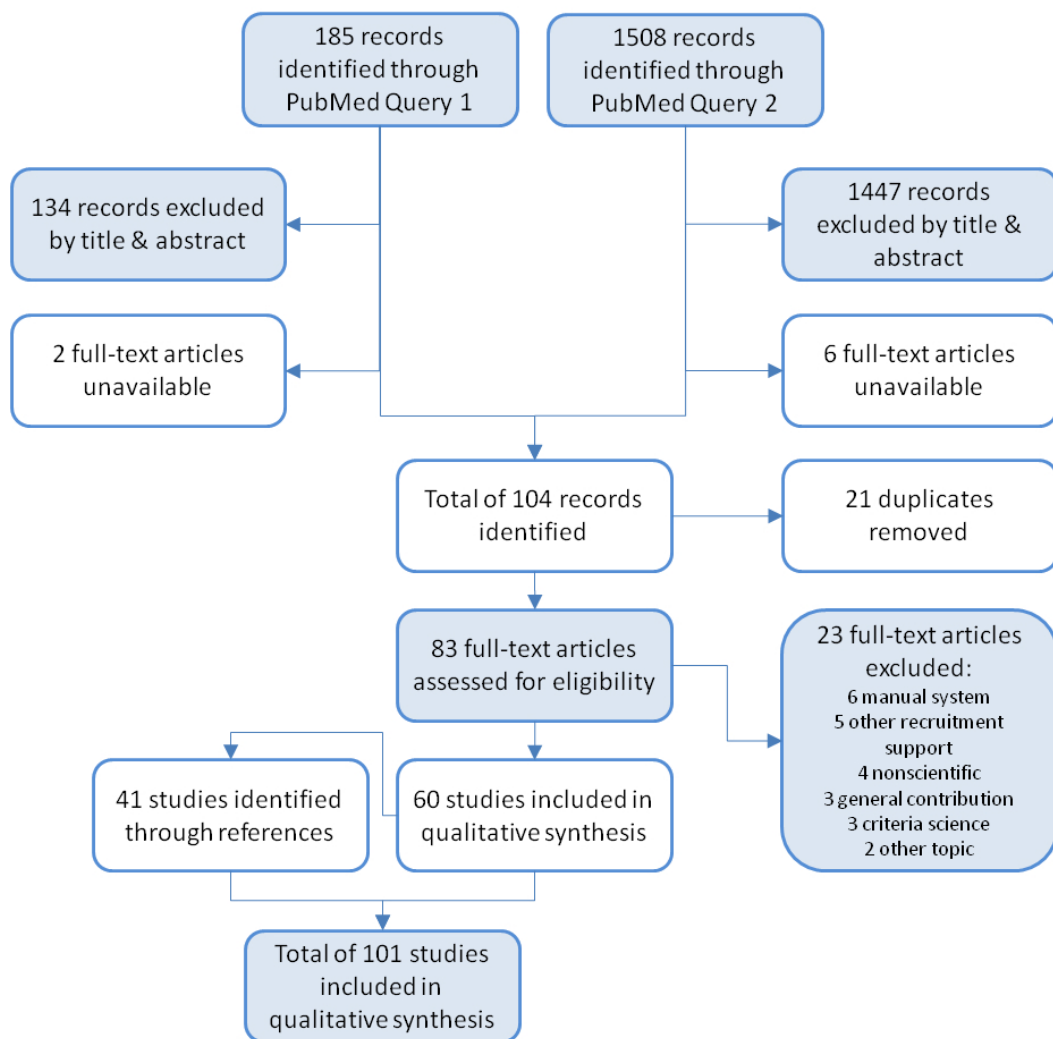
Results Structure

Multimedia Appendix 1 shows a list of all articles grouped by system and ordered by first publication date (objective 1). It also summarizes the CTRSS characteristics according to the categories described subsequently. In the following sections on CTRSS characteristics, we identify and describe CTRSS groups with similar features (objective 2). We also speculate on environmental characteristics that led the developers to favor a group or reject another (objective 3). All evidence for the impact of CTRSS on patient recruitment is presented in Outcomes (objective 4).

Characteristics of Included Articles

Regarding system maturity, 12 articles reported on a CTRSS concept that was not implemented yet. A total of 42 articles described a prototypical implementation, often including performance tests, but no application to a running clinical trial. Another 47 articles described fully matured systems that were used to recruit patients into at least 1 trial. First publications on CTRSS dated back to 1990. However, there were no more than 3 publications per year until 2003. Since then, 7 articles per year were published on average, so that nearly 80% of all articles were from the past 10 years (Figure 2).

Figure 2. Number of publications on CTRSS per year.
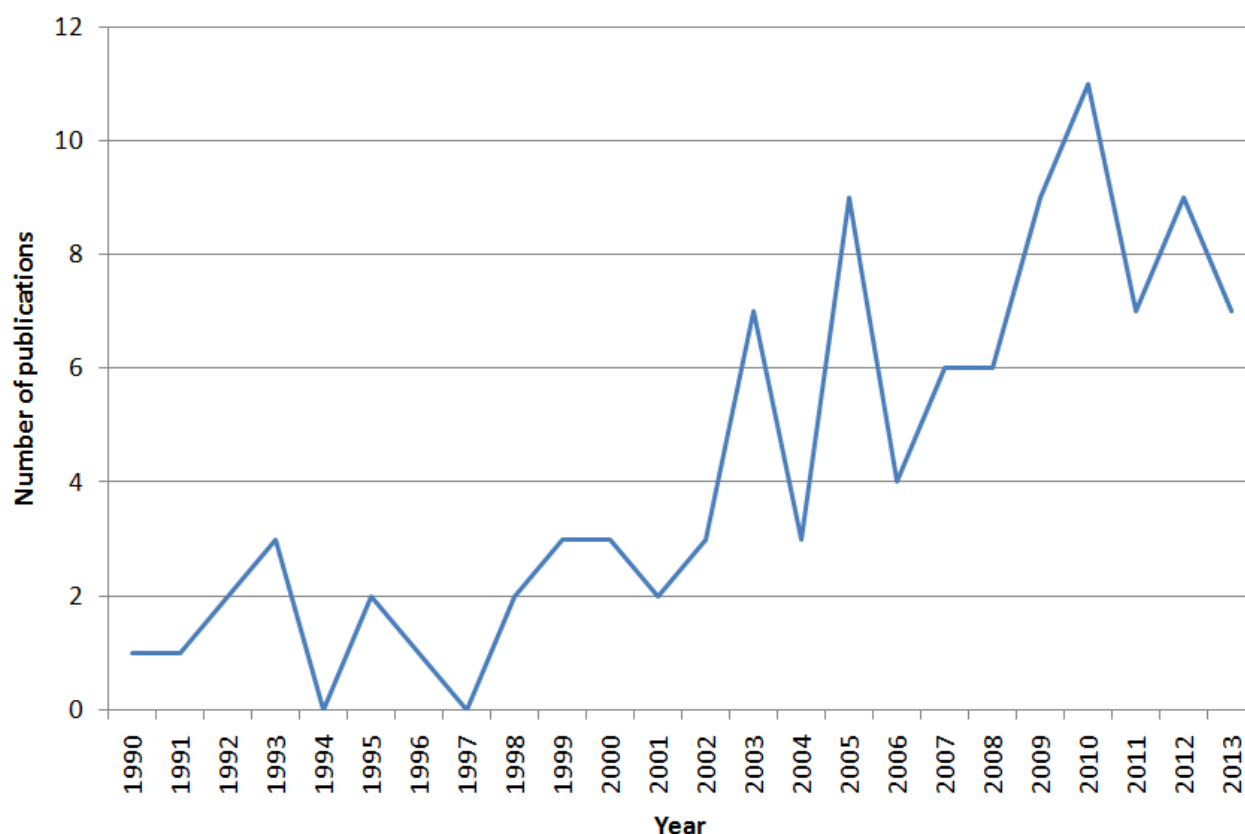
Clinical Context and Scope of Application

CTRSS have been implemented and used in trials in a wide variety of clinical domains. Still, many systems were evaluated for only 1 trial or trials from the same domain. With 17 CTRSS in this domain, oncology (especially breast cancer) was found particularly often. This domain may be favorable because it is research intensive with many open trials and exceptionally large available volumes of patient data and funding. The functionalities and algorithms of the CTRSS seemed largely independent from the clinical domain. Thus, no author precluded the use of their system for clinical trials from other domains and many actually suggested it.

The accuracy of a CTRSS depends on the available patient data and its effect depends on the organizational environment in which it operates. Therefore, each CTRSS should be evaluated for a large number of trials and at multiple sites to increase the reliability of reported results if possible. Many authors observed

this: 43 articles reported on using their system for more than 1 trial (11 did not name an exact figure) and 14 CTRSS were intended for use at multiple sites. In comparison, 37 reports evaluated a CTRSS for a single trial and 62 CTRSS were used at a single institution. In all, 11 papers failed to give the number of trials their CTRSS had been evaluated or used for.

Workflow Integration

Overview

Every CTRSS has 2 points of contact with the recruitment workflow of a clinical trial. The first is the trigger that causes the system to assess the eligibility of one or more patients. The second is the communication of the assessment's results (eg, a list of potential trial participants) to the system's user.

Trigger

One way to trigger the eligibility assessment is to have the user or an administrator execute a manual process. Manual triggers

are both the easiest to implement and the most commonly found. They are sufficient for cases in which patient data are entered into the CTRSS by the user who can subsequently view patient eligibility in an interactive fashion. The user can be a physician [38,88] or the patient [59,79,112]. Manual triggers are also sufficient for cases in which an eligibility assessment is required only once to generate a patient list, which is not expected to change during the trial's recruitment phase. The latter is generally the case for retrospective trials and feasibility studies. Typical examples include Payne et al [97], Thadani et al [115], and Köpcke et al [69] who required an administrator to develop a Structured Query Language (SQL)-based query. Based on 16 years of COSTAR research queries, Murphy [85] created the graphical interface Informatics for Integrating Biology and the Bedside (i2b2) to allow investigators to parameterize query templates themselves.

For trials that require regular re-evaluation of patient eligibility because of changing patient data over time, manual triggers are generally inefficient and are replaced by automatic triggers. Automatic triggers can start eligibility assessments periodically at given time intervals [40,122] or in reaction to particular events in the hospital information system (HIS) [28,44,48,60]. Time-based triggers are generally easier to implement than event-based triggers. The interval length between assessments depends on the requirements of each trial and the computing time required for an assessment. It is usually set to a value between several minutes and 1 day. For trials that require an immediate reaction to new patient data by trial staff and for trials with comparatively rare potential participants, trigger events are preferred. Such triggers include the availability of new data or the admission of a patient.

Communication

The results of an eligibility assessment must be communicated somehow to the CTRSS user. The primary factor of influence when choosing a mode of communication is the target user group. If patients are supposed to use the CTRSS, it is most common to offer a separate user interface that interactively displays potentially fitting trials and/or a score indicating the patient's fit with a certain trial [32,59,72,79,109,112,125]. Exceptions are found if the patient is interested in future trials instead of ones that are currently recruiting. In these cases, patients enter their health data into a registry or a personal health record and they are notified by email as soon as a fitting trial is detected [61,125]. If the CTRSS has no clinical/research user (ie, the direct user is IT staff), it usually transforms the raw result of the reasoning algorithm into a patient list which is subsequently handed out to the researcher [60,80,103,105,115,119]. This is the preferred mode of communication if eligibility assessments are only required once [69,97]. However, when the target users are either treating physicians or clinical investigators, the mode of communication also needs to accommodate data security regulations and the trial's temporal requirements. Pagers seem to be the only option if the user needs to react immediately to new patient data, such as critical laboratory values [41,63]. When time is of less importance, emails are chosen to deliver both proposals for single patients and patient sets alike [34,45,106,124]. A recurring scenario is that the physician or nurse is reminded of a trial during their first consultation. To achieve this, alerts or flags are placed in the EHR which appear at a convenient moment and often allow direct evaluation of the patient's eligibility [104,121].

When coupled with simultaneous messages, automatic triggers have the disadvantage of easily initiating alerts or prompts at a time when the user is not prepared to answer them. Untimely messages will cause the receiver to ignore them. The same effect occurs for systems with a large share of false positive alerts. This alert fatigue is regularly mentioned as a problem for CTRSS efficiency and acceptance. Numbers for the fraction of alerts that are actually reviewed by the receiver range from 25% [60], more than 30% to 40% [52], and 56% [105] to less than 70% [62]. For Ruffin [105], even "numerous prompts and reminders and customized requests" could not solve the problem. Additionally, Embi and Leonard [52] found that response rates declined at a rate of 2.7% per 2-week time period.

Knowledge Representation and Data Sources

Overview

The core technical functionality of a CTRSS is the comparison of eligibility criteria with the electronically available patient data. According to Weng et al [124], the process is characterized by 3 aspects: "the expression language for representing eligibility rules, the encoding of eligibility concepts, and the modeling of patient data." The underlying problem is that eligibility criteria are almost always given in narrative form and need to be translated into a structure that can be interpreted by the CTRSS. The same is true for the patient data, which needs to be analyzed to identify concepts that match the eligibility concepts before developing the eligibility rules themselves.

Source of Patient Data

Most authors choose the data source for their CTRSS according to availability and accessibility. Few CTRSS designs are based on a comparison of different potential data sources (eg, for timeliness or comprehensiveness). Nevertheless, the reuse of existing patient data for the purpose of recruitment is common practice: 64 CTRSS relied on data that was collected for other purposes originally. A total of 5 monitored the health level 7 (HL7) messages of a clinical information system, 46 of them read patient data directly from the EHR of the hospital or general practitioner, 12 used a data warehouse, and 1 used a clinical registry. In this order, these data sources increasingly collect and integrate patient data over time, software applications, and institutions, which makes access to the data of large patient sets comparatively easy. However, more integrated data often means the data source becomes increasingly detached from its origin as well (ie, some information is lost during processing and delays between the documented event and availability of the corresponding data grow). For some trials, such delays are unacceptable because trial staff need to be notified as soon as possible for specific events. Specialty subsystems, such as an electronic tracking board [3,29] or the messages exchanged between these systems [54,76,98,101,121], need to be monitored directly in these cases. A total of 3 CTRSS preloaded patient characteristics from the EHR and prompted the physician to complete missing data [90,94,95]. Wilcox et al [125]

conceptualized a CTRSS that integrated EHR data and the personal health record of a patient. Only 16 CTRSS made exclusive use of data that were entered directly into the system itself by the physician (n=8), the patient (n=7), or an investigator (n=1).

Terminologies

The CTRSS developer can choose the terminology for clinical concept names arbitrarily if patient data are entered only for the purpose of eligibility assessment. However, if patient data are taken from an already existing data source, most developers chose to reuse the terminologies found there. A total of 66 articles did not mention the use of any terminology. Of these, 5 performed pure free-text analysis and did not necessarily require terminologies. Of those papers that did mention the use of a specific terminology, 16 named the International Classification of Diseases (ICD). This makes sense because it is also the terminology most commonly used within EHRs. There were no other widespread terminologies used for CTRSS. The Unified Medical Language System (UMLS) appeared in 6 publications and the Medical Entities Dictionary (MED), Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), and Logical Observation Identifiers Names and Codes (LOINC) in only 3, respectively. In all, 10 terminologies were each used in only 1 CTRSS, such as Cerner Multum, National Drug Code, Hospital International Classification of Diseases Adapted, Read, NCI-Thesaurus (NCI-T), and 837 billing data, or in 2 CTRSS, such as MeSH, NCI Common Data Elements (CDE), and Current Procedural Terminology.

Intermediary Criteria Format

Terminologies are usually chosen to suit the available patient data, whereas the intermediary criteria format is strongly associated with the reasoning method of the CTRSS. The SQL is the most frequently found representation of criteria logic. Unfortunately, the CTRSS literature lacks details on the representation of criteria expressions. A comparison of the eligibility criteria as given in the study protocol and their representation in the CTRSS is rare; 49 papers gave no information on the chosen format of eligibility expressions.

Translation Process

With a few exceptions, the translation process to make eligibility criteria processable for the computer seemed to be a manual one. For 51 CTRSS, the administrator was responsible for reading the trial protocol, mapping clinical concepts to the target terminology, and creating eligibility expressions. This is the most efficient process in clinical settings that generate few trials per researcher because teaching costs are minimized and experience is concentrated in 1 person. Yet, a notable fraction of the CTRSS offered the user an interface to select eligibility criteria autonomously from a small [38] or large [43] set of predefined criteria. Having the user translate the eligibility criteria of a trial is primarily meaningful for feasibility studies, giving a researcher the means to dynamically modify the criteria for a new trial and to instantly receive feedback for the change's influence on the expected number of participants.

Lonsdale et al [75] proposed natural language processing (NLP) to support the translation process. They read eligibility sentences from the trial registry ClinicalTrials.gov, parsed them to retrieve logical forms and mapped concepts to standard terminologies to generate executable Arden syntax Medical Logic Modules (MLMs). The process succeeded for 16% of all criteria from 85 randomly chosen trials [74,75]. Zhang et al [127] and Köpcke et al [70] proposed case-based reasoning algorithms for free-text and structured patient data, respectively. These algorithms did not require the translation of eligibility criteria into rules, but tried to determine the unknown eligibility of new patients by comparing them with a set of patients with known eligibility status.

Reasoning

Overview

Closely tied to the previously described CTRSS characteristics is the reasoning process itself (ie, the method to assess whether the available data for a patient suffices for the conditions set by the trial's eligibility criteria). Almost all CTRSS "perform 'pre-screening' for clinical research staff" [115] instead of trying to determine the actual eligibility of a patient. They do not replace manual chart review, but act as a filter that limits the number of patients who require such by selecting the most likely candidates. The presentation of reasoning details, such as a probability of eligibility or missing patient characteristics together with the screening list, can facilitate the manual screening process even further.

The dominance of relational databases for the storage of patient data entails that most CTRSS employ database queries somewhere in the reasoning process. Consequentially, most CTRSS are based on an elaborate query or a set of subsequently executed queries per trial [3,45,64,126]. If the result set of potentially eligible patients is sufficiently accurate, no further processing is required.

Some authors demonstrated the feasibility of more exotic reasoning methods. A total of 4 CTRSS used Arden syntax to control the reasoning process [64,75,91,95]; 3 CTRSS employed an ontologic reasoner after transforming eligibility criteria and, in 2 cases, patient data into separate ontologies [35,72,96]. However, although technically interesting, the authors failed to convey the advantages of these algorithms when compared with the aforementioned simpler ones.

Dealing With Incomplete Data

Some CTRSS designers paid particular attention to missing patient data. Tu et al [119] developed 2 methods for dealing with this problem. In their qualitative method, each criterion was attributed 1 of 5 qualities according to a patient's concrete data: patient meets the criterion, patient probably meets the criterion, no assertion possible, patient probably fails the criterion, and patient fails the criterion. Specific rules for each criterion derived one of these qualities from the patient's data or assign default values. In their probabilistic method, a Bayesian belief network was manually constructed for each trial. The network represented variables as nodes and dependencies as links between nodes. All nodes and links were given probabilities based on legacy data or experts. If data for

a variable were found, the variable was given a probability of 1 or zero; otherwise, the default probabilities were used. When all available data for a patient were retrieved, a probability for the patient's eligibility could be calculated. This probabilistic approach was applied again later by Papaconstantinou et al [94] and Ash et al [31]. Bhanja et al [36] suggested that scalability as well as time and design complexities discouraged the use of probabilistic approaches.

Natural Language Processing

The wish to include unstructured (ie, free-text) data could also warrant the utilization of complex reasoning algorithms. Keyword searches were often employed when no structured data elements were available [29,41,73,93,101,106,110]. They could easily be added to complement queries of structured patient data [3,66,98,124]. Pakhomov [93] compared a keyword search with 2 other NLP methods: naive Bayes and perceptron. Naive Bayes yielded the best sensitivity (95% vs 86% and 71% for perceptron and keyword search, respectively) and perceptron offered the best specificity (65% vs 57% and 54% for naive Bayes and keyword search, respectively). Although performing worst of all methods, the advantage of using a simple keyword search lies in its easy implementation (no need for training data)

and transparency. In a similar comparison, Zhang et al [127] found regular expressions outperformed a vector space method and latent semantic indexing to achieve accuracy similar to a specifically developed method called subtree match. However, they also proposed algorithms for automatic keyword and subtree generation, which could offer distinct potential for automation.

Sensitivity-Versus-Specificity Tradeoff

Independent from the chosen reasoning, the inclusiveness of each CTRSS is subject to the desires of its user. Ultimately, the setup of a CTRSS "requires sensitivity-versus-specificity tradeoffs" for each trial [119]. The upper limit to specificity might be determined by the fit between available patient data and eligibility criteria, whereas its lower limit is simply determined by what the user is willing to accept (Figure 3). The required level of sensitivity is limited by the availability of trial participants. Sensitivity should be chosen as low as possible to increase specificity and, thus, reduce recruitment workload. In practice, however, when the CTRSS is motivated by a lack of participants for a specific trial, maximum sensitivity is imperative and low specificity must be accepted [49].

Figure 3. The sensitivity and specificity of patient proposals from most CTRSS depends on the configuration of the reasoning algorithm. The developer is free to favor specificity over sensitivity or vice versa, depending on conditions that are likely to be different for each trial. Frequently relevant conditions are user acceptance, the availability of patient data, and the availability of participants compared to the required number.
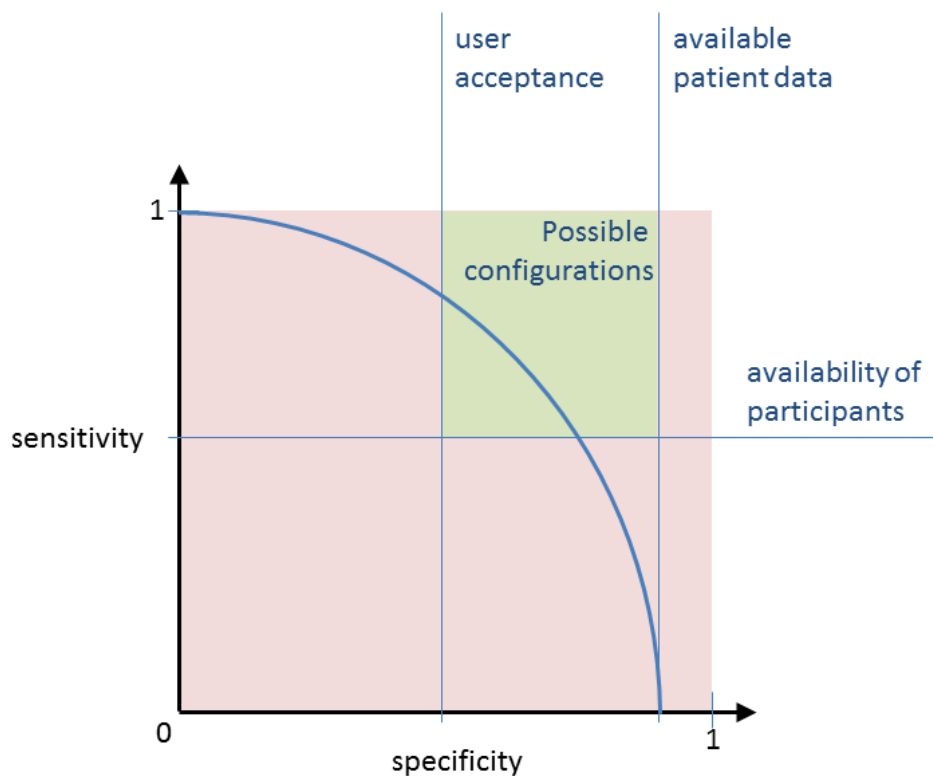
Outcome

Overview

All studies in this review shared the common goal to improve the recruitment process of clinical trials. However, calculating the performance of the CTRSS in terms of specificity and sensitivity alone is, at best, a secondary indicator for its effect. Direct comparison with the manual recruitment process with regard to its effects on one or more of the following 3 variables should be favored: (1) the pure number of trial participants (ie, the effectiveness of the recruitment process), (2) the cost to recruit a given number of patients in terms of money and/or time (ie, the efficiency of the recruitment process), and (3) the quality of the collective of trial participants (eg, measures for selection bias and dropouts). All reported system effects were weighted according to the scientific quality of the evaluation as (1) reliable quantitative measurement, (2) quantitative measurement with insufficient description of or flawed method, or (3) survey or estimation (corresponding to A-C in Multimedia Appendix 1, respectively).

Impact on Recruitment Effectiveness

We found 5 papers that reliably quantified differences in recruitment effectiveness between manual and CTRSS-supported recruitment. Embi et al [49] reported on a doubling of physician's enrollment rate from 3 to 6 per month, which was attributed to a concurrent significant increase in the number of referring physicians from 5 to 42. The CTRSS presented by Cardozo et al [41] increased identification of eligible patients from 1 in 2 months to 6 in 2 months after physicians failed to generate pager notifications in time. Herasevich et al [63] doubled monthly enrollment rates from 37 in approximately 8.5 months to 68 in approximately 9 months in a time-critical setting. They attributed the effect to the change from imprecise clinical notes (manual process) to specific physiologic criteria (automated process) as the basis for eligibility evaluation. Beauharnais et al [34] also doubled recruitment, in this case from 11 patients in 63 days to 20 patients in 62 days. The effect seemed to correlate with an increase in screening efficiency that similarly doubled the number of screened patients. A comparatively minor increase in recruited patients of 14% from 306 to 348 in the same week was reported by Köpcke et al [69] who addressed pure oversight of otherwise well-organized manual recruiters. They also found 7% of the manually included patients did not fulfill the trial's eligibility criteria.

Lane et al [71], Tu et al [119], and a research group from the University of South Florida [55,56,68,88] ran their respective CTRSS on legacy patient data and evaluated how many of those patients found potentially eligible by their system were actually enrolled in the past. These works only showed an upper limit of CTRSS effectiveness because it was unclear whether "physicians actually missed the matches, rather than having undocumented reasons for omitting them" [56]. Similarly, Weiner et al [122] described an increase in the number of eligibility alerts sent to the trial investigator. Again, these can only be an upper limit for the effect of the CTRSS on enrollment because the physician's reasons for not alerting the investigator were unclear. It is possible that the physicians judged patients unfit for the trial for reasons beyond the criteria that were

considered by the CTRSS or that the patients were unwilling to participate. Séroussi and Bouaud [108], Weng et al [124], and Treweek et al [118] compared the effectiveness of their CTRSS with conventional methods of recruitment by running them in parallel over the whole study period. However, the lack of enrollment numbers for a preceding phase without the CTRSS made it impossible to quantify the effect of the CTRSS. Finally, Ferranti et al [54] reported an increase in recruitment numbers by 53%. Although we found their methodology suitable, the authors failed to discuss reasons for a sharp increase in recruitment numbers 2 months before introduction of the CTRSS.

Impact on Recruitment Efficiency

We judged 4 papers to reliably quantify differences in the efficiency of a CTRSS and the manual recruitment process. Thompson et al [117] reduced the screening time required per eligible patient from 18 to 6 minutes (66%) in a 2-week evaluation of their CTRSS prototype. This reduction was achieved solely through a higher fraction of eligible patients among screened patients, whereas the individual screening time was actually higher for patients proposed by the CTRSS. Penberthy et al [98] verified this circumstance for 5 additional trials, achieving screening time reductions of 95%, 34%, 86%, and 34% in 4 trials and an increase of 31% in 1 trial. Again, time savings resulted from screening fewer noneligible patients, whereas individual screening time remained unchanged. Therefore, the benefit in efficiency was found to depend on the specificity of the CTRSS. Nkoy et al [3] decreased screening time from 2 to zero hours daily with no manual control of the patient list generated by their CTRSS. They translated these time savings into cost savings of US $1200 per month. Beauharnais et al [34] halved screening time from 4 to 2 hours daily, measuring manual and CTRSS-aided recruitment over 60 subsequent days, respectively. They concluded that "the use of an algorithm is most beneficial for studies with low enrollment rates because of the long duration of the accrual period."

Following a proposition by Ohno-Machado et al [90], the aforementioned research group from the University of South Florida [55,56,68,88] presented a unique approach to increase screening efficiency. Through ordering of the necessary clinical tests for eligibility determination in such a way that cheap but decisive tests were done first, they expected a reduction of costs by 50%. The cost of each test and the number of clinical trials and eligibility criteria that required a test's results were included in the calculation. Unfortunately, the evaluation of the methodology was based on retrospective data and it remained unclear how the cost for tests without reordering were calculated. Seyfried et al [110] reported decreased screening time, but used the same dataset with the same test physicians for both manual and CTRSS-aided screenings (50 patients, 1-week interval). Furthermore, the CTRSS appeared to be trained with the same dataset on which it was tested later. Thadani et al [115] and Schmickl et al [106] did not directly measure screening time decreases, but stated that they could imagine screening only patients proposed by their respective CTRSS to be sufficient, reducing the patient pool by 81% and

76%, respectively. Obviously, such a strategy would require the CTRSS to feature a sufficiently high sensitivity.

Impact on Recruitment Quality

Only Rollman et al [104] compared the characteristics of patient sets after manual and CTRSS-aided recruitment. To this end, they observed 2 subsequent trials with similar eligibility criteria, the same recruitment period of 22 months and the same 4 recruiting primary care physicians. They found that usage of the CTRSS significantly increased the proportion of male nonwhite patients, as well as the fraction of patients with more severe disease grades.

Discussion

Principal Findings

There are some CTRSS setups that reappear on a regular basis. Firstly, for the retrospective identification of trial participants based on existing clinical data, database queries are designed and executed once or on a regular basis. They create a list of potentially eligible patients that is printed on paper or otherwise delivered to the researcher. Secondly, for trials with short windows of opportunity for recruitment, a key event in the EHR or another health IT component is constantly monitored. Its occurrence causes a more comprehensive eligibility test for the concerned patient and is communicated to the researcher via pager. Thirdly, if no patient data exist yet, it is entered directly into the CTRSS, which assesses and communicates the patient's eligibility directly after completion of data entry. Although the treating physician was required to act for the patient in older systems, it is now becoming increasingly popular to offer this possibility directly to the patient via dedicated websites. Our review confirms the findings of Weng et al [129] who also gave names to these CTRSS types: (1) mass screening decision support, (2) EHR-based recruitment alerts, and (3) computerized research protocol systems and Web-based patient-enabling systems (depending on the user).

The setup of a specific CTRSS is rarely chosen on a theoretical background (ie, after an evaluation of different options for triggering the system and communicating the results). Instead, the setup is dictated mostly by the existing clinical environment, available IT tools, and the needs of a specific trial or group of researchers. Because CTRSS are a subset of clinical decision support systems (CDSS), it will generally be possible to configure existing CDSS such that they assume CTRSS functionalities (eg, [50,64,86]).

Limitations of the Review

Our review is limited in that the collection of publications and extraction of information from these publications was done by only 1 author. We reduced the impact of this approach by refraining from any interpretation of the given information in this step. Nevertheless, we cannot preclude mistakes, especially when stating that no or unclear information on a certain CTRSS characteristic was found in an article. Furthermore, all unreferenced statements made in this review reflect only the opinion of the 2 authors and are subject to discussion by the research community.

Comparison With Previous Review

Our review of 101 CTRSS publications offers the most comprehensive and up-to-date overview on CTRSS. Compared to the previous review paper by Cuggia et al from 2011 [17], which analyzed 28 CTRSS from articles published before October 2009, we identified an increase of publications in the subsequent years. These more recent publications present more data on the impact of CTRSS on the recruitment process, which we discuss subsequently. Of the 7 tendencies in CTRSS research formulated by Cuggia et al, all but the exclusive reliance on structured data appear to continue. We found many CTRSS that include unstructured data as a data source, although many of them are limited to keyword searches. There are 3 additional lessons we believe can be learned from the existing research, which are described subsequently.

Lack of standards is not limited to the terminologies of the patient data source, but also applies to the computational representation of eligibility criteria. Although researchers have proposed independent languages to encode the free-text criteria of a trial's protocol (eg, ERGO [130], EliXR [131]), most CTRSS bind the representation of eligibility criteria in 2 ways to the specifics of their environment: (1) to the terminology of the patient data source and (2) to the chosen reasoning method. We believe independent and exchangeable eligibility criteria to be desirable because multisite trials have become the norm. However, judging from the experience so far, readily encoded criteria will need to be the norm in trial protocols before they will be adopted by CTRSS designers. Tools to help translate the criteria into SQL statements could speed up the adoption process.

The choice of the reasoning method should consider its pervasiveness (ie, how easily third parties interested in its deployment can learn to install and administrate it). Considering this, no other method seems to be as suitable for CTRSS as SQL queries on relational databases. Queries can make use of existing data from the EHR, a data warehouse (DWH), or a registry and their administrators are likely to be experienced creators and users of such queries. Resistance to adopt and maintain an additional query-based system is likely to be small compared to CTRSS that require additional training in one of the less widespread technologies, such as probabilistic methods or Arden syntax. Although complex reasoning methods have been shown to achieve high accuracy, it is unclear whether they lead to an increased CTRSS impact compared to queries.

Using patient care data promises efficiency and effectiveness gains for a CTRSS. But, because it is collected for other purposes, it also introduces new challenges [131]. It is imperfect from the viewpoint of eligibility assessments because it lacks uniformity (the same information can be documented differently for 2 patients), timeliness (information might be documented too late), and completeness (information might be missing for some or all patients). Uniformity and completeness problems can lead to severe selection bias and increase the cost of eligibility rule creation. For example, low uniformity necessitates an analysis of documentation habits; low completeness might enforce the use of proxy data [78] or estimates [90]. Timeliness must be ensured by the

documentation process, which might resist change. Untimely data will severely limit the possibility to support a trial, especially in outpatient settings [46]. Thus, unfit data can constitute a major limitation to CTRSS impact.

First Conclusions on Clinical Trial Recruitment Support Systems Impact

We suggested that the introduction of a CTRSS can be motivated by 3 expectations: (1) an increase in the number of participants for a given clinical trial or a set of trials, (2) a reduction of trial costs through decreased screening costs, and (3) the guarantee to select a representative set of patients (ie, the reduction of selection bias). Many authors do not elaborate on the shortcomings of the manual recruitment process that led to the development of their CTRSS.

Whether a CTRSS is able to increase the number of participants for a trial depends little on its setup, but rather on the deficits of the manual recruitment process it is set to replace. To begin with, an untapped group of potential participants (ie, a gap between those patients who are eligible and those who are asked to participate) needs to exist. This gap originates from some patients not being screened at all or from communication problems between the different actors of the recruitment process. Thus, a CTRSS can close this gap if it can ensure that every patient is screened and that the necessary information on the patient and the trial is available in time.

Often, a CTRSS is expected to close the gap between estimated and realized participant numbers or that between eligible and recruited patients or even the gap between needed and available patients. These expectations are likely to be disappointed. They disregard that many causes for insufficient recruitment are out of the scope of a CTRSS or simply cannot be addressed by an IT intervention. The most important is a willingness of the patient to participate and motivation of physicians to participate in recruitment. The analysis of the existing recruitment process and its weaknesses should, therefore, be part of every CTRSS design process. Weng et al [123,124] give examples on how to do this. They characterize patient eligibility status in different categories, such as potentially eligible, approachable, consentable, eligible, and ultimately enrolled. By comparing the ratio of patients in each category, such taxonomy can be used to identify the weak spots in recruitment that need to be addressed by the CTRSS.

Although the effectiveness of a CTRSS is determined by its setting, improvements in screening efficiency might be more generally achievable. Many successes to reduce screening time

are based on using existing data to reliably exclude patients from the screening list (ie, the CTRSS generates no or few false negatives). In this way, the CTRSS can be used to reduce the number of patients that must be screened manually. Under the reasonable assumption that documented data are correct, but not all patient characteristics are documented, we believe CTRSS should focus on the exclusion criteria of a clinical trial to maximize efficiency gains. No final eligibility decision should be based on the trial's inclusion criteria because this can reduce the sensitivity of the CTRSS and motivate the screeners to use other screening methods in parallel. To realize efficiency gains, the CTRSS must completely replace the former screening process. This also means that the aim to increase recruitment efficiency is opposed to the other 2 potential aims of a CTRSS which profit from running multiple screening methods in parallel.

The potential benefit of a CTRSS on the composition of a trial's participants has been insufficiently explored so far. Because patient demographics should be easily obtainable for all experiments comparing manual and CTRSS-aided recruitment, we suggest including them in future publications.

Future Directions

We found most articles describe the characteristics and operating principles of their CTRSS reasonably well, but all lacked in some regard. Intermediary criteria representation, terminologies of the patient data, and an evaluation of the system's effects were often missing. Many authors present prototypes of their CTRSS directly after finishing its design and fail to report on its outcome and usage. We encourage more follow-up publications on the experiences with existing CTRSS such as those by Embi et al [51], Embi and Leonard [52], and Dugas et al [47]. To strengthen the comprehensibility and usefulness of future reports, we propose a list of essential elements that should be included (Textbox 1).

In their review of patient cohort identification systems in general, Shivade et al [21] found a "growing trend in the areas of machine learning and data mining" and believe these necessary to develop generalizable solutions. For CTRSS in particular, this trend has not yet manifested in the literature. Only Zhang et al [127] and Köpcke et al [70] report on experiments to exploit these techniques for recruitment purposes, but both are still in a prototype stage. Machine learning promises more independence from the individual representation of patient data in a hospital and better portability. Still more data are needed to assess advantages and disadvantages and to explore hybrid solutions.

Textbox 1. Essential elements to be included in future CTRSS studies.

Clinical Context

1.   Number of trials that the CTRSS has been evaluated for

2.   Length of time the CTRSS has been in use

3.   Brief description and number of sites that use the CTRSS

Input

1.   Representation format of eligibility criteria in the CTRSS

2.   Comparison of original and computable eligibility criteria for an exemplary trial

3.   Summary of how well the eligibility criteria could be translated for all other trials (if any)

4.   Details on the translation process

5.   Representation of patient data in the CTRSS

6.   Details on the patient data source (eg, purpose, terminologies)

Working Principle

1.   What triggers an eligibility assessment?

2.   How are eligibility criteria and patient data compared?

3.   How is the result of the assessment communicated, when is it communicated, and to whom?

Outcome

1.   Recruitment process before introduction of CTRSS including perceived problems

2.   Recruitment process following introduction of CTRSS

3.   Patient numbers and time spent for each step

Current publications in the area of CTRSS are still too focused on—and sometimes limited to—technical aspects of system setup and the accuracy of its eligibility assessment. After review of most of the existing literature, we believe that the impact of a CTRSS on a given recruitment process is determined more by the context of the CTRSS (ie, the available patient data, its integration in trial, and clinical workflows and its attraction to users). Therefore, what is needed are research projects to evaluate how a CTRSS can be embedded in different recruitment workflows, the characteristics of trials that profit from CTRSS, different designs for user interaction, and the outcomes of CTRSS in relation to these parameters.

Conclusions

We further found that differences in the setup of CTRSS are because of existing infrastructure and particularities of the recruitment process, particularly the target user of the CTRSS (eg, treating physician, study nurse) and the prior recruitment problem (eg, failure to identify, failure to communicate). Yet, there are still many questions open in defining when and how CTRSS can best improve recruitment processes in clinical trials. Based on the questions that remained open in our analysis of many of the 101 articles, we propose an item list that should be considered for future publications on CTRSS design, implementation, and evaluation. This shall ensure that CTRSS setup and background, their integration in research processes, and their outcome results are sufficiently described to allow researchers to better learn from other´s experiences.

Authors' Contributions

FK conducted the review, wrote, and reviewed the manuscript. HUP wrote and reviewed the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Table of publications included in qualitative analysis together with major CTRSS characteristics.

[XLSX File (Microsoft Excel File), 31KB - jmir_v16i7e161_app1.xlsx]

References

1. Tassignon JP, Sinackevich N. Speeding the Critical Path. Appl Clin Trials 2004 Jan;13(1):42 [FREE Full text]
2. Sullivan J. Subject recruitment and retention: Barriers to success. Appl Clin Trials 2004 Apr:50-54 [FREE Full text]
3. Nkoy FL, Wolfe D, Hales JW, Lattin G, Rackham M, Maloney CG. Enhancing an existing clinical information system to improve study recruitment and census gathering efficiency. AMIA Annu Symp Proc 2009;2009:476-480 [FREE Full text] [Medline: 20351902]
4. Hunninghake DB, Darby CA, Probstfield JL. Recruitment experience in clinical trials: literature summary and annotated bibliography. Control Clin Trials 1987 Dec;8(4 Suppl):6S-30S. [Medline: 3326716]
5. Gotay CC. Accrual to cancer clinical trials: directions from the research literature. Soc Sci Med 1991;33(5):569-577. [Medline: 1962228]
6. Ellis PM. Attitudes towards and participation in randomised clinical trials in oncology: a review of the literature. Ann Oncol 2000 Aug;11(8):939-945 [FREE Full text] [Medline: 11038029]
7. Collins JF, Williford WO, Weiss DG, Bingham SF, Klett CJ. Planning patient recruitment: fantasy and reality. Stat Med 1984 Dec;3(4):435-443. [Medline: 6396795]
8. Campbell MK, Snowdon C, Francis D, Elbourne D, McDonald AM, Knight R, STEPS group. Recruitment to randomised trials: strategies for trial enrollment and participation study. The STEPS study. Health Technol Assess 2007 Nov;11(48):iii, ix-iii,105 [FREE Full text] [Medline: 17999843]
9. Taylor KM, Kelner M. Interpreting physician participation in randomized clinical trials: the Physician Orientation Profile. J Health Soc Behav 1987 Dec;28(4):389-400. [Medline: 3429808]
10. Siminoff LA, Zhang A, Colabianchi N, Sturm CM, Shen Q. Factors that predict the referral of breast cancer patients onto clinical trials by their surgeons and medical oncologists. J Clin Oncol 2000 Mar;18(6):1203-1211. [Medline: 10715289]
11. Mannel RS, Walker JL, Gould N, Scribner DR, Kamelle S, Tillmanns T, et al. Impact of individual physicians on enrollment of patients into clinical trials. Am J Clin Oncol 2003 Apr;26(2):171-173. [doi: 10.1097/01.COC.0000017798.43288.7C] [Medline: 12714890]
12. Rahman M, Morita S, Fukui T, Sakamoto J. Physicians' reasons for not entering their patients in a randomized controlled trial in Japan. Tohoku J Exp Med 2004 Jun;203(2):105-109 [FREE Full text] [Medline: 15212145]
13. Comis RL, Miller JD, Aldigé CR, Krebs L, Stoval E. Public attitudes toward participation in cancer clinical trials. J Clin Oncol 2003 Mar 1;21(5):830-835. [Medline: 12610181]
14. Grunfeld E, Zitzelsberger L, Coristine M, Aspelund F. Barriers and facilitators to enrollment in cancer clinical trials: qualitative study of the perspectives of clinical research associates. Cancer 2002 Oct 1;95(7):1577-1583 [FREE Full text] [doi: 10.1002/cncr.10862] [Medline: 12237928]
15. McDonald AM, Knight RC, Campbell MK, Entwistle VA, Grant AM, Cook JA, et al. What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. Trials 2006;7:9 [FREE Full text] [doi: 10.1186/1745-6215-7-9] [Medline: 16603070]
16. Toerien M, Brookes ST, Metcalfe C, de Salis I, Tomlin Z, Peters TJ, et al. A review of reporting of participant recruitment and retention in RCTs in six major journals. Trials 2009;10:52 [FREE Full text] [doi: 10.1186/1745-6215-10-52] [Medline: 19591685]
17. Cuggia M, Besana P, Glasspool D. Comparing semi-automatic systems for recruitment of patients to clinical trials. Int J Med Inform 2011 Jun;80(6):371-388. [doi: 10.1016/j.ijmedinf.2011.02.003] [Medline: 21459664]
18. Coorevits P, Sundgren M, Klein GO, Bahr A, Claerhout B, Daniel C, et al. Electronic health records: new opportunities for clinical research. J Intern Med 2013 Dec;274(6):547-560. [doi: 10.1111/joim.12119] [Medline: 23952476]
19. Trinczek B, Köpcke F, Leusch T, Majeed RW, Schreiweis B, Wenk J, et al. Design and multicentric implementation of a generic software architecture for patient recruitment systems re-using existing HIS tools and routine patient data. Appl Clin Inform 2014;5(1):264-283 [FREE Full text] [doi: 10.4338/ACI-2013-07-RA-0047] [Medline: 24734138]
20. Hripcsak G, Knirsch C, Zhou L, Wilcox A, Melton G. Bias associated with mining electronic health records. J Biomed Discov Collab 2011;6:48-52 [FREE Full text] [doi: 10.5210/disco.v6i0.3581] [Medline: 21647858]
21. Shivade C, Raghavan P, Fosler-Lussier E, Embi PJ, Elhadad N, Johnson SB, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc 2014;21(2):221-230 [FREE Full text] [doi: 10.1136/amiajnl-2013-001935] [Medline: 24201027]
22. Andersen MR, Schroeder T, Gaul M, Moinpour C, Urban N. Using a population-based cancer registry for recruitment of newly diagnosed patients with ovarian cancer. Am J Clin Oncol 2005 Feb;28(1):17-20. [Medline: 15685029]
23. Ashburn A, Pickering RM, Fazakarley L, Ballinger C, McLellan DL, Fitton C. Recruitment to a clinical trial from the databases of specialists in Parkinson's disease. Parkinsonism Relat Disord 2007 Feb;13(1):35-39. [doi: 10.1016/j.parkreldis.2006.06.004] [Medline: 16928464]

24. Tu SW, Peleg M, Carini S, Bobak M, Ross J, Rubin D, et al. A practical method for transforming free-text eligibility criteria into computable criteria. J Biomed Inform 2011 Apr;44(2):239-250 [FREE Full text] [doi: 10.1016/j.jbi.2010.09.007] [Medline: 20851207]

25. Patel CO, Weng C. ECRL: an eligibility criteria representation language based on the UMLS Semantic Network. AMIA Annu Symp Proc 2008:1084. [Medline: 18999200]

26. Friedlin J, Overhage M, Al-Haddad MA, Waters JA, Aguilar-Saavedra JJ, Kesterson J, et al. Comparing methods for identifying pancreatic cancer patients using electronic data sources. AMIA Annu Symp Proc 2010;2010:237-241 [FREE Full text] [Medline: 21346976]

27. Carroll RJ, Eyler AE, Denny JC. Naïve Electronic Health Record phenotype identification for Rheumatoid arthritis. AMIA Annu Symp Proc 2011;2011:189-196 [FREE Full text] [Medline: 22195070]

28. Afrin LB, Oates JC, Boyd CK, Daniels MS. Leveraging of open EMR architecture for clinical trial accrual. AMIA Annu Symp Proc 2003:16-20 [FREE Full text] [Medline: 14728125]

29. Ahmad F, Gupta R, Kurz M. Real time electronic patient study enrollment system in emergency room. AMIA Annu Symp Proc 2005:881 [FREE Full text] [Medline: 16779168]

30. Ainsworth J, Buchan I. Preserving consent-for-consent with feasibility-assessment and recruitment in clinical studies: FARSITE architecture. Stud Health Technol Inform 2009;147:137-148. [Medline: 19593052]

31. Ash N, Ogunyemi O, Zeng Q, Ohno-Machado L. Finding appropriate clinical trials: evaluating encoded eligibility criteria with incomplete data. Proc AMIA Symp 2001:27-31 [FREE Full text] [Medline: 11825151]

32. Atkinson NL, Massett HA, Mylks C, Hanna B, Deering MJ, Hesse BW. User-centered research on breast cancer patient needs and preferences of an Internet-based clinical trial matching system. J Med Internet Res 2007;9(2):e13 [FREE Full text] [doi: 10.2196/jmir.9.2.e13] [Medline: 17513284]

33. Bache R, Miles S, Taweel A. An adaptable architecture for patient cohort identification from diverse data sources. J Am Med Inform Assoc 2013 Dec;20(e2):e327-e333. [doi: 10.1136/amiajnl-2013-001858] [Medline: 24064442]

34. Beauharnais CC, Larkin ME, Zai AH, Boykin EC, Luttrell J, Wexler DJ. Efficacy and cost-effectiveness of an automated screening algorithm in an inpatient clinical trial. Clin Trials 2012 Apr;9(2):198-203. [doi: 10.1177/1740774511434844] [Medline: 22308560]

35. Besana P, Cuggia M, Zekri O, Bourde A, Burgun A. Using semantic web technologies for clinical trial recruitment. : Springer; 2010 Presented at: The Semantic Web-ISWC 2010: 9th International Semantic Web Conference, ISWC 2010; November 7-11, 2010; Shanghai, China p. 7748.

36. Bhanja S, Fletcher-Heath LM, Hall LO, Goldgof DB, Krischer JP. A qualitative expert system for clinical trial assignment. In: FLAIRS-98: Proceedings of the Eleventh International Florida Artificial Intelligence Research Symposium Conference.: AAAI Press; 1998 Presented at: Eleventh International Florida Artificial Intelligence Research Symposium Conference; May 18-20, 1998; Sanibel Island, Florida p. 84.

37. Boland MR, Miotto R, Gao J, Weng C. Feasibility of feature-based indexing, clustering, and search of clinical trials. A case study of breast cancer trials from ClinicalTrials.gov. Methods Inf Med 2013;52(5):382-394 [FREE Full text] [doi: 10.3414/ME12-01-0092] [Medline: 23666475]

38. Breitfeld PP, Weisburd M, Overhage JM, Sledge G, Tierney WM. Pilot study of a point-of-use decision support tool for cancer clinical trials eligibility. J Am Med Inform Assoc 1999 Dec;6(6):466-477 [FREE Full text] [Medline: 10579605]

39. Breitfeld PP, Ullrich F, Anderson J, Crist WM. Web-based decision support for clinical trial eligibility determination in an international clinical trials network. Control Clin Trials 2003 Dec;24(6):702-710. [Medline: 14662275]

40. Butte AJ, Weinstein DA, Kohane IS. Enrolling patients into clinical trials faster using RealTime Recuiting. Proc AMIA Symp 2000:111-115 [FREE Full text] [Medline: 11079855]

41. Cardozo E, Meurer WJ, Smith BL, Holschen JC. Utility of an automated notification system for recruitment of research subjects. Emerg Med J 2010 Oct;27(10):786-787. [doi: 10.1136/emj.2009.081299] [Medline: 20515907]

42. Carlson RW, Tu SW, Lane NM, Lai TL, Kemper CA, Musen MA, et al. Computer-based screening of patients with HIV/AIDS for clinical-trial eligibility. Online J Curr Clin Trials 1995 Mar 28;Doc No 179:[3347 words; 32 paragraphs]. [Medline: 7719564]

43. Deshmukh VG, Meystre SM, Mitchell JA. Evaluating the informatics for integrating biology and the bedside system for clinical research. BMC Med Res Methodol 2009;9:70 [FREE Full text] [doi: 10.1186/1471-2288-9-70] [Medline: 19863809]

44. Duftschmid G, Gall W, Eigenbauer E, Dorda W. Management of data from clinical trials using the ArchiMed system. Med Inform Internet Med 2002 Jun;27(2):85-98. [doi: 10.1080/1463923021000014158] [Medline: 12507268]

45. Dugas M, Lange M, Berdel WE, Müller-Tidow C. Workflow to improve patient recruitment for clinical trials within hospital information systems - a case-study. Trials 2008;9:2 [FREE Full text] [doi: 10.1186/1745-6215-9-2] [Medline: 18186949]

46. Dugas M, Amler S, Lange M, Gerss J, Breil B, Köpcke W. Estimation of patient accrual rates in clinical trials based on routine data from hospital information systems. Methods Inf Med 2009;48(3):263-266. [doi: 10.3414/ME0582] [Medline: 19387510]

47. Dugas M, Lange M, Müller-Tidow C, Kirchhof P, Prokosch HU. Routine data from hospital information systems can support patient recruitment for clinical studies. Clin Trials 2010 Apr;7(2):183-189. [doi: 10.1177/1740774510363013] [Medline: 20338903]

XSL•FO
RenderX

48.   Embi PJ, Jain A, Harris CM. Physician perceptions of an Electronic Health Record-based Clinical Trial Alert system: a survey of study participants. AMIA Annu Symp Proc 2005:949 [FREE Full text] [Medline: 16779236]

49.   Embi PJ, Jain A, Clark J, Bizjack S, Hornung R, Harris CM. Effect of a clinical trial alert system on physician participation in trial recruitment. Arch Intern Med 2005 Oct 24;165(19):2272-2277 [FREE Full text] [doi: 10.1001/archinte.165.19.2272] [Medline: 16246994]

50.   Embi PJ, Jain A, Clark J, Harris CM. Development of an electronic health record-based Clinical Trial Alert system to enhance recruitment at the point of care. AMIA Annu Symp Proc 2005:231-235 [FREE Full text] [Medline: 16779036]

51.   Embi PJ, Jain A, Harris CM. Physicians' perceptions of an electronic health record-based clinical trial alert approach to subject recruitment: a survey. BMC Med Inform Decis Mak 2008;8:13 [FREE Full text] [doi: 10.1186/1472-6947-8-13] [Medline: 18384682]

52.   Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. J Am Med Inform Assoc 2012 Jun;19(e1):e145-e148 [FREE Full text] [doi: 10.1136/amiajnl-2011-000743] [Medline: 22534081]

53.   Fernández-Breis JT, Maldonado JA, Marcos M, Legaz-García Mdel C, Moner D, Torres-Sospedra J, et al. Leveraging electronic healthcare record standards and semantic web technologies for the identification of patient cohorts. J Am Med Inform Assoc 2013 Dec;20(e2):e288-e296. [doi: 10.1136/amiajnl-2013-001923] [Medline: 23934950]

54.   Ferranti JM, Gilbert W, McCall J, Shang H, Barros T, Horvath MM. The design and implementation of an open-source, data-driven cohort recruitment system: the Duke Integrated Subject Cohort and Enrollment Research Network (DISCERN). J Am Med Inform Assoc 2012 Jun;19(e1):e68-e75 [FREE Full text] [doi: 10.1136/amiajnl-2011-000115] [Medline: 21946237]

55.   Fink E, Hall LO, Goldgof DB, Goswami BD, Boonstra M, Krischer JP. Experiments on the automated selection of patients for clinical trials. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. 2003 Presented at: IEEE International Conference on Systems, Man, and Cybernetics; October 5-8, 2003; Washington, DC p. 4541-4545. [doi: 10.1109/ICSMC.2003.1245699]

56.   Fink E, Kokku PK, Nikiforou S, Hall LO, Goldgof DB, Krischer JP. Selection of patients for clinical trials: an interactive web-based system. Artif Intell Med 2004 Jul;31(3):241-254. [doi: 10.1016/j.artmed.2004.01.017] [Medline: 15302090]

57.   Gennari JH, Reddy M. Participatory design and an eligibility screening tool. Proc AMIA Symp 2000:290-294 [FREE Full text] [Medline: 11079891]

58.   Gennari JH, Sklar D, Silva J. Cross-tool communication: from protocol authoring to eligibility determination. Proc AMIA Symp 2001:199-203 [FREE Full text] [Medline: 11825180]

59.   Graham AL, Cha S, Cobb NK, Fang Y, Niaura RS, Mushro A. Impact of seasonality on recruitment, retention, adherence, and outcomes in a web-based smoking cessation intervention: randomized controlled trial. J Med Internet Res 2013;15(11):e249 [FREE Full text] [doi: 10.2196/jmir.2880] [Medline: 24201304]

60.   Grundmeier RW, Swietlik M, Bell LM. Research subject enrollment by primary care pediatricians using an electronic health record. AMIA Annu Symp Proc 2007:289-293 [FREE Full text] [Medline: 18693844]

61.   Harris PA, Lane L, Biaggioni I. Clinical research subject recruitment: the Volunteer for Vanderbilt Research Program www.volunteer.mc.vanderbilt.edu. J Am Med Inform Assoc 2005 Dec;12(6):608-613 [FREE Full text] [doi: 10.1197/jamia.M1722] [Medline: 16049231]

62.   Heinemann S, Thüring S, Wedeken S, Schäfer T, Scheidt-Nave C, Ketterer M, et al. A clinical trial alert tool to recruit large patient samples and assess selection bias in general practice research. BMC Med Res Methodol 2011;11:16 [FREE Full text] [doi: 10.1186/1471-2288-11-16] [Medline: 21320358]

63.   Herasevich V, Pieper MS, Pulido J, Gajic O. Enrollment into a time sensitive clinical study in the critical care setting: results from computerized septic shock sniffer implementation. J Am Med Inform Assoc 2011 Oct;18(5):639-644 [FREE Full text] [doi: 10.1136/amiajnl-2011-000228] [Medline: 21508415]

64.   Jenders RA, Hripcsak G, Sideli RV, DuMouchel W, Zhang H, Cimino JJ, et al. Medical decision support: experience with implementing the Arden Syntax at the Columbia-Presbyterian Medical Center. Proc Annu Symp Comput Appl Med Care 1995:169-173 [FREE Full text] [Medline: 8563259]

65.   Kamal J, Pasuparthi K, Rogers P, Buskirk J, Mekhjian H. Using an information warehouse to screen patients for clinical trials: a prototype. AMIA Annu Symp Proc 2005:1004 [FREE Full text] [Medline: 16779291]

66.   Kho A, Zafar A, Tierney W. Information technology in PBRNs: the Indiana University Medical Group Research Network (IUMG ResNet) experience. J Am Board Fam Med 2007 Apr;20(2):196-203 [FREE Full text] [doi: 10.3122/jabfm.2007.02.060114] [Medline: 17341757]

67.   Koca M, Husmann G, Jesgarz J, Overath M, Brandts C, Serve H. A special query tool in the hospital information system to recognize patients and to increase patient numbers for clinical trials. Stud Health Technol Inform 2012;180:1180-1181. [Medline: 22874393]

68.   Kokku PK, Hall LO, Goldgof DB, Fink E, Krischer JP. A cost-effective agent for clinical trial assignment. In: Proceedings of the 2002 IEEE International Conference on Systems, Man & Cybernetics: bridging the digital divide: cyber-development, human progress, peace and prosperity. 2002 Presented at: 2002 IEEE International Conference on Systems, Man & Cybernetics; October 6-9, 2002; Yasmine Hammamet, Tunisia.

69. Köpcke F, Kraus S, Scholler A, Nau C, Schüttler J, Prokosch HU, et al. Secondary use of routinely collected patient data in a clinical trial: an evaluation of the effects on patient recruitment and data acquisition. Int J Med Inform 2013 Mar;82(3):185-192. [doi: 10.1016/j.ijmedinf.2012.11.008] [Medline: 23266063]

70. Köpcke F, Lubgan D, Fietkau R, Scholler A, Nau C, Stürzl M, et al. Evaluating predictive modeling algorithms to assess patient eligibility for clinical trials from routine data. BMC Med Inform Decis Mak 2013;13:134 [FREE Full text] [doi: 10.1186/1472-6947-13-134] [Medline: 24321610]

71. Lane NM, Kemper CA, Fodor M, Tu SW, Lai T, Yvon M, et al. Opportunity to enhance accrual to clinical trials using a microcomputer. 1991 Presented at: International Conference On AIDS: Science Challenging AIDS; June 16-21, 1991; Florence, Italy p. 16-21.

72. Lee Y, Dinakarpandian D, Katakam N, Owens D. MindTrial: An Intelligent System for Clinical Trials. AMIA Annu Symp Proc 2010;2010:442-446 [FREE Full text] [Medline: 21347017]

73. Li L, Chase HS, Patel CO, Friedman C, Weng C. Comparing ICD9-encoded diagnoses and NLP-processed discharge summaries for clinical trials pre-screening: a case study. AMIA Annu Symp Proc 2008:404-408 [FREE Full text] [Medline: 18999285]

74. Lonsdale D, Tustison C, Parker C, Embley DW. Formulating queries for assessing clinical trial eligibility. In: Proceedings of the 11th international conference on Applications of Natural Language to Information Systems. Berlin: Springer; 2006 Presented at: NLDB'06: 11th international conference on Applications of Natural Language to Information Systems; May 31-June 2, 2006; Klagenfurt, Austria p. 82-93. [doi: 10.1007/11765448_8]

75. Lonsdale DW, Tustison C, Parker CG, Embley DW. Assessing clinical trial eligibility with logic expression queries. Data & Knowledge Engineering 2008 Jul;66(1). [doi: 10.1016/j.datak.2007.07.005]

76. Majeed RW, Röhrig R. Identifying patients for clinical trials using fuzzy ternary logic expressions on HL7 messages. Stud Health Technol Inform 2011;169:170-174. [Medline: 21893736]

77. Marcos M, Maldonado JA, Martínez-Salvador B, Boscá D, Robles M. Interoperability of clinical decision-support systems and electronic health records using archetypes: a case study in clinical trial eligibility. J Biomed Inform 2013 Aug;46(4):676-689 [FREE Full text] [doi: 10.1016/j.jbi.2013.05.004] [Medline: 23707417]

78. McGregor J, Brooks C, Chalasani P, Chukwuma J, Hutchings H, Lyons RA, et al. The Health Informatics Trial Enhancement Project (HITE): Using routinely collected primary care data to identify potential participants for a depression trial. Trials 2010;11:39 [FREE Full text] [doi: 10.1186/1745-6215-11-39] [Medline: 20398303]

79. Metz JM, Coyle C, Hudson C, Hampshire M. An Internet-based cancer clinical trials matching resource. J Med Internet Res 2005 Jul 1;7(3):e24 [FREE Full text] [doi: 10.2196/jmir.7.3.e24] [Medline: 15998615]

80. Miller JL. The EHR solution to clinical trial recruitment in physician groups. Health Manag Technol 2006 Dec;27(12):22-25. [Medline: 17256648]

81. Miller DM, Fox R, Atreja A, Moore S, Lee JC, Fu AZ, et al. Using an automated recruitment process to generate an unbiased study sample of multiple sclerosis patients. Telemed J E Health 2010 Feb;16(1):63-68 [FREE Full text] [doi: 10.1089/tmj.2009.0078] [Medline: 20064056]

82. Mosis G, Koes B, Dieleman J, Stricker B, van der Lei J, Sturkenboom MC. Randomised studies in general practice: how to integrate the electronic patient record. Inform Prim Care 2005;13(3):209-213. [Medline: 16259861]

83. Mosis G, Dieleman JP, Stricker B, van der Lei J, Sturkenboom MC. A randomized database study in general practice yielded quality data but patient recruitment in routine consultation was not practical. J Clin Epidemiol 2006 May;59(5):497-502. [doi: 10.1016/j.jclinepi.2005.11.007] [Medline: 16632138]

84. Murphy SN, Morgan MM, Barnett GO, Chueh HC. Optimizing healthcare research data warehouse design through past COSTAR query analysis. Proc AMIA Symp 1999:892-896 [FREE Full text] [Medline: 10566489]

85. Murphy SN, Barnett GO, Chueh HC. Visual query tool for finding patient cohorts from a clinical data warehouse of the partners HealthCare system. Proc AMIA Symp 2000:1174 [FREE Full text] [Medline: 11080028]

86. Musen MA, Carlson RW, Fagan LM, Deresinski SC, Shortliffe EH. T-HELPER: automated support for community-based clinical research. Proc Annu Symp Comput Appl Med Care 1992:719-723 [FREE Full text] [Medline: 1482965]

87. Nalichowski R, Keogh D, Chueh HC, Murphy SN. Calculating the benefits of a Research Patient Data Repository. AMIA Annu Symp Proc 2006:1044 [FREE Full text] [Medline: 17238663]

88. Kokku PK, Hall LO, Goldgof DB, Fink E, Krischer JP. A cost-effective agent for clinical trial assignment. In: Proceedings of the 2002 IEEE International Conference on Systems, Man & Cybernetics: bridging the digital divide: cyber-development, human progress, peace and prosperity. Piscataway, N.J: IEEE; 2002 Presented at: 2002 IEEE International Conference on Systems, Man & Cybernetics: bridging the digital divide: cyber-development, human progress, peace and prosperity; October 6-9, 2002; Yasmine Hammamet, Tunisia.

89. Oberg R, Rasmussen L, Melski J, Peissig P, Starren J. Evaluation of the google search appliance for patient cohort discovery. AMIA Annu Symp Proc 2008:1104. [Medline: 18999072]

90. Ohno-Machado L, Parra E, Henry SB, Tu SW, Musen MA. AIDS2: a decision-support tool for decreasing physicians' uncertainty regarding patient eligibility for HIV treatment protocols. Proc Annu Symp Comput Appl Med Care 1993:429-433 [FREE Full text] [Medline: 8130510]

91. Ohno-Machado L, Wang SJ, Mar P, Boxwala AA. Decision support for clinical trial eligibility determination in breast cancer. Proc AMIA Symp 1999:340-344 [FREE Full text] [Medline: 10566377]

92. Olesen F, Kjeldsen HC, Christensen MB. Identifying patients for research in general practice. Scand J Prim Health Care 1996 Mar;14(1):62-63. [Medline: 8725096]

93. Pakhomov SV, Buntrock J, Chute CG. Prospective recruitment of patients with congestive heart failure using an ad-hoc binary classifier. J Biomed Inform 2005 Apr;38(2):145-153 [FREE Full text] [doi: 10.1016/j.jbi.2004.11.016] [Medline: 15797003]

94. Papaconstantinou C, Theocharous G, Mahadevan S. An expert system for assigning patients into clinical trials based on Bayesian networks. J Med Syst 1998 Jun;22(3):189-202. [Medline: 9604785]

95. Parker CG, Embley DW. Generating medical logic modules for clinical trial eligibility criteria. AMIA Annu Symp Proc 2003:964 [FREE Full text] [Medline: 14728468]

96. Patel C, Cimino JJ, Dolby J, Fokoue A, Kalyanpur A, Kershenbaum A, et al. Matching patient records to clinical trials using ontologies. In: Proceedings of the 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007.: Springer; 2007 Presented at: The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007; November 11-15, 2007; Busan, Korea p. 6297.

97. Payne TH, Goroll AH, Morgan M, Barnett GO. Conducting a matched-pairs historical cohort study with a computer-based ambulatory medical record system. Comput Biomed Res 1990 Oct;23(5):455-472. [Medline: 2225790]

98. Penberthy L, Brown R, Puma F, Dahman B. Automated matching software for clinical trials eligibility: measuring efficiency and flexibility. Contemp Clin Trials 2010 May;31(3):207-217. [doi: 10.1016/j.cct.2010.03.005] [Medline: 20230913]

99. Peterson KA, Delaney BC, Arvanitis TN, Taweel A, Sandberg EA, Speedie S, et al. A model for the electronic support of practice-based research networks. Ann Fam Med 2012 Dec;10(6):560-567 [FREE Full text] [doi: 10.1370/afm.1434] [Medline: 23149534]

100. Pressler TR, Yen PY, Ding J, Liu J, Embi PJ, Payne PR. Computational challenges and human factors influencing the design and use of clinical research participant eligibility pre-screening tools. BMC Med Inform Decis Mak 2012;12:47 [FREE Full text] [doi: 10.1186/1472-6947-12-47] [Medline: 22646313]

101. Quinn J, Durski K. A real-time tracking, notification, and web-based enrollment system for emergency department research. Acad Emerg Med 2004 Nov;11(11):1245-1248. [doi: 10.1197/j.aem.2004.08.020] [Medline: 15528591]

102. 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]

103. Rokach L, Maimon O, Averbuch M. Information retrieval system for medical narrative reports. In: Flexible Query Answering Systems: 6th International Conference, FQAS 2004, Proceedings (Lecture Notes in Computer Science / Lecture Notes in Artificial Intelligence).: Springer; 2004 Presented at: Flexible Query Answering Systems: 6th International Conference, FQAS 2004; June 24-26, 2004; Lyon, France p. 217.

104. Rollman BL, Fischer GS, Zhu F, Belnap BH. Comparison of electronic physician prompts versus waitroom case-finding on clinical trial enrollment. J Gen Intern Med 2008 Apr;23(4):447-450 [FREE Full text] [doi: 10.1007/s11606-007-0449-0] [Medline: 18373143]

105. Ruffin MT, Nease DE. Using patient monetary incentives and electronically derived patient lists to recruit patients to a clinical trial. J Am Board Fam Med 2011 Oct;24(5):569-575 [FREE Full text] [doi: 10.3122/jabfm.2011.05.100169] [Medline: 21900440]

106. Schmickl CN, Li M, Li G, Wetzstein MM, Herasevich V, Gajic O, et al. The accuracy and efficiency of electronic screening for recruitment into a clinical trial on COPD. Respir Med 2011 Oct;105(10):1501-1506 [FREE Full text] [doi: 10.1016/j.rmed.2011.04.012] [Medline: 21576012]

107. Schreiweis B, Grüber G, Bergh B. First experiences in implementing a software-based support for patient recruitment at Heidelberg university hospital. Stud Health Technol Inform 2012;180:1147-1149. [Medline: 22874382]

108. Séroussi B, Bouaud J. Using OncoDoc as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials. Artif Intell Med 2003 Oct;29(1-2):153-167. [Medline: 12957785]

109. Seward J, Follansbee S, Christianson B. Development of a computer database system to assist patients with enrollment into an HIV clinical trial. 1992 Presented at: VII International Conference on AIDS and the II STD World Congress; July 19-24, 1992; Amsterdam, Netherlands.

110. Seyfried L, Hanauer DA, Nease D, Albeiruti R, Kavanagh J, Kales HC. Enhanced identification of eligibility for depression research using an electronic medical record search engine. Int J Med Inform 2009 Dec;78(12):e13-e18 [FREE Full text] [doi: 10.1016/j.ijmedinf.2009.05.002] [Medline: 19560962]

111. Silva MI, Abbott P, Petrucci K, Canfield K, Muir J. Improving the efficiency of patient recruitment with an automated telephone screening system in a client-server environment. Proc Annu Symp Comput Appl Med Care 1993:41-45 [FREE Full text] [Medline: 8130506]

112. Smith KS, Eubanks D, Petrik A, Stevens VJ. Using web-based screening to enhance efficiency of HMO clinical trial recruitment in women aged forty and older. Clin Trials 2007;4(1):102-105. [doi: 10.1177/1740774506075863] [Medline: 17327250]

113. Stell A, Sinnott R, Ajayi O. Supporting the clinical trial recruitment process through the grid. In: Proceedings of the UK e-Science All Hands Meeting 2006. Edinburgh: National e-Science Centre; 2006 Presented at: UK e-Science All Hands Meeting; September 18-21, 2006; Nottingham p. 61-68.

114. Sumi E, Teramukai S, Yamamoto K, Satoh M, Yamanaka K, Yokode M. The correlation between the number of eligible patients in routine clinical practice and the low recruitment level in clinical trials: a retrospective study using electronic medical records. Trials 2013;14:426 [FREE Full text] [doi: 10.1186/1745-6215-14-426] [Medline: 24326039]

115. Thadani SR, Weng C, Bigger JT, Ennever JF, Wajngurt D. Electronic screening improves efficiency in clinical trial recruitment. J Am Med Inform Assoc 2009 Dec;16(6):869-873 [FREE Full text] [doi: 10.1197/jamia.M3119] [Medline: 19717797]

116. Thew S, Leeming G, Ainsworth J, Gibson M, Buchan I. FARSITE: evaluation of an automated trial feasibility assessment and recruitment tool. Trials 2011;12(Suppl 1):A113. [doi: 10.1186/1745-6215-12-S1-A113]

117. Thompson DS, Oberteuffer R, Dorman T. Sepsis alert and diagnostic system: integrating clinical systems to enhance study coordinator efficiency. Comput Inform Nurs 2003 Feb;21(1):22-6; quiz 27. [Medline: 12544151]

118. Treweek S, Pearson E, Smith N, Neville R, Sargeant P, Boswell B, et al. Desktop software to identify patients eligible for recruitment into a clinical trial: using SARMA to recruit to the ROAD feasibility trial. Inform Prim Care 2010;18(1):51-58. [Medline: 20429978]

119. Tu SW, Kemper CA, Lane NM, Carlson RW, Musen MA. A methodology for determining patients' eligibility for clinical trials. Methods Inf Med 1993 Aug;32(4):317-325. [Medline: 8412828]

120. Vahabzadeh M, Lin JL, Mezghanni M, Contoreggi C, Leff M. A clinical recruiting management system for complex multi-site clinical trials using qualification decision support systems. AMIA Annu Symp Proc 2007:1141. [Medline: 18694237]

121. Weber S, Lowe HJ, Malunjkar S, Quinn J. Implementing a Real-time Complex Event Stream Processing System to Help Identify Potential Participants in Clinical and Translational Research Studies. AMIA Annu Symp Proc 2010;2010:472-476 [FREE Full text] [Medline: 21347023]

122. Weiner DL, Butte AJ, Hibberd PL, Fleisher GR. Computerized recruiting for clinical trials in real time. Ann Emerg Med 2003 Feb;41(2):242-246. [doi: 10.1067/mem.2003.52] [Medline: 12548275]

123. Weng C, Bigger JT, Busacca L, Wilcox A, Getaneh A. Comparing the effectiveness of a clinical registry and a clinical data warehouse for supporting clinical trial recruitment: a case study. AMIA Annu Symp Proc 2010;2010:867-871 [FREE Full text] [Medline: 21347102]

124. Weng C, Batres C, Borda T, Weiskopf NG, Wilcox AB, Bigger JT, et al. A real-time screening alert improves patient recruitment efficiency. AMIA Annu Symp Proc 2011;2011:1489-1498 [FREE Full text] [Medline: 22195213]

125. Wilcox A, Natarajan K, Weng C. Using Personal Health Records for Automated Clinical Trials Recruitment: the ePaIRing Model. Summit on Translat Bioinforma 2009;2009:136-140 [FREE Full text] [Medline: 21347187]

126. Yamamoto K, Sumi E, Yamazaki T, Asai K, Yamori M, Teramukai S, et al. A pragmatic method for electronic medical record-based observational studies: developing an electronic medical records retrieval system for clinical research. BMJ Open 2012;2(6) [FREE Full text] [doi: 10.1136/bmjopen-2012-001622] [Medline: 23117567]

127. Zhang J, Gu Y, Liu W, Hu W, Zhao T, Mu X, et al. Automatic patient search for breast cancer clinical trials using free-text medical reports. In: Proceedings of ACM International Health Informatics Symposium, IHI 2010. 2010 Presented at: IHI '10 Proceedings of the 1st ACM International Health Informatics Symposium; November 11-12, 2010; Arlington, VA p. 405-409. [doi: 10.1145/1882992.1883052]

128. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]

129. Weng C, Tu SW, Sim I, Richesson R. Formal representation of eligibility criteria: a literature review. J Biomed Inform 2010 Jun;43(3):451-467 [FREE Full text] [doi: 10.1016/j.jbi.2009.12.004] [Medline: 20034594]

130. Tu SW, Peleg M, Carini S, Bobak M, Ross J, Rubin D, et al. A practical method for transforming free-text eligibility criteria into computable criteria. J Biomed Inform 2011 Apr;44(2):239-250 [FREE Full text] [doi: 10.1016/j.jbi.2010.09.007] [Medline: 20851207]

131. Weng C, Wu X, Luo Z, Boland MR, Theodoratos D, Johnson SB. EliXR: an approach to eligibility criteria extraction and representation. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i116-i124 [FREE Full text] [doi: 10.1136/amiajnl-2011-000321] [Medline: 21807647]

132. Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PR, Bernstam EV, et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. Med Care 2013 Aug;51(8 Suppl 3):S30-S37. [doi: 10.1097/MLR.0b013e31829b1dbd] [Medline: 23774517]

Abbreviations

CDE: common data elements
CTRSS: clinical trial recruitment support systems

EHR: electronic health record
ERGO: Eligibility Rule Grammar and Ontology
HIS: hospital information system
ICD: International Classification of Diseases
LOINC: Logical Observation Identifiers Names and Codes
MED: Medical Entities Dictionary
MLM: Medical Logic Module
NCI: National Cancer Institute
NLP: natural language processing
SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms
SQL: Structured Query Language
UMLS: Unified Medical Language System

XSL•FO
RenderX

Original Paper

Does Self-Selection Affect Samples' Representativeness in Online Surveys? An Investigation in Online Video Game Research

Yasser Khazaal[1], MD; Mathias van Singer[2]; Anne Chatton[1], MA; Sophia Achab[1], MD, PhD; Daniele Zullino[2], MD; Stephane Rothen[2], PhD; Riaz Khan[1], MD; Joel Billieux[3], PhD; Gabriel Thorens[1], MD

[1]Geneva University Hospitals, Geneva, Switzerland

[2]Geneva University, Geneva, Switzerland

[3]Laboratory for Experimental Psychopathology, Psychological Science Research Institute, Catholic University of Louvain, Louvain-La-Neuve, Belgium

Corresponding Author:
Yasser Khazaal, MD
Geneva University Hospitals
Grand pré, 70 C
Geneva, 1206
Switzerland
Phone: 41 795535682
Fax: 41 223202840
Email: yasser.khazaal@hcuge.ch

Abstract

Background: The number of medical studies performed through online surveys has increased dramatically in recent years. Despite their numerous advantages (eg, sample size, facilitated access to individuals presenting stigmatizing issues), selection bias may exist in online surveys. However, evidence on the representativeness of self-selected samples in online studies is patchy.

Objective: Our objective was to explore the representativeness of a self-selected sample of online gamers using online players' virtual characters (avatars).

Methods: All avatars belonged to individuals playing World of Warcraft (WoW), currently the most widely used online game. Avatars' characteristics were defined using various games' scores, reported on the WoW's official website, and two self-selected samples from previous studies were compared with a randomly selected sample of avatars.

Results: We used scores linked to 1240 avatars (762 from the self-selected samples and 478 from the random sample). The two self-selected samples of avatars had higher scores on most of the assessed variables (except for guild membership and exploration). Furthermore, some guilds were overrepresented in the self-selected samples.

Conclusions: Our results suggest that more proficient players or players more involved in the game may be more likely to participate in online surveys. Caution is needed in the interpretation of studies based on online surveys that used a self-selection recruitment procedure. Epidemiological evidence on the reduced representativeness of sample of online surveys is warranted.

Introduction

An increasing number of medical and psychological studies are performed through online surveys. Compared with face-to-face interviews, Internet-based surveys can quickly reach more potential participants, reduce measurement error and bias related to answers on stigmatizing topics, and enhance the inclusion of least represented or "quasi-secret" and stigmatized population groups that are usually difficult to reach and recruit [1-5]. Costs can be more easily contained with Internet-based surveys, and

data collection can be simpler and more reliable compared to traditional paper-and-pencil data entry procedures. Some studies suggest that the quality of the data provided by Internet-based surveys is at least as good as in those collected by traditional paper-and-pencil methods on self-selected samples [6-8].

Some Web surveys have been based on the assessment of a whole population or on samples obtained using random sampling procedures (ie, a sampling technique whereby all individuals in the population have an equal chance of being selected, eg, emailing a random sample of students in a university). For

XSL•FO
RenderX

instance, Internet-based surveys among students enrolled by email have generated valid and reliable estimations of substance use [3,9,10], comparable to those obtained in studies that applied ordinary mail invitation letters or phone calls to recruit participants.

Many Web studies are, however, self-selection surveys [11] that are not based on probability sampling [12], particularly in health-related studies. Websites and online social networks such as Facebook appeared to be a viable recruitment option for the assessment of health behaviors [13,14]. However, lack of researchers' knowledge about the website members' contacts leads to the impossibility of obtaining a random sampling. The survey questionnaire is then usually put on the Web. Potential participants are among those people with Internet access who visit the website, find the study information, and decide to complete the survey. In the case of self-selection surveys, the researcher then has no control over the selection process and can work only on the design of the study advertisement (such as graphics and content, length of questions, possible incentives) or on a selection of an appropriate website or forum to promote the response rate to electronic questionnaires [15]. Online self-selection surveys are thus particularly subject to coverage and selection bias, which undermines the external validity of studies and the interpretation of findings [12].

Coverage bias is possibly influenced by patterns related to Internet access or to specific website access (ie, differences between people with or without Internet access) and to the possibility of being particularly interested in the study for reasons that may or may not be related to the content and/or objective of the survey itself. Furthermore, exposure to the advertisement is influenced by the time spent on a specific website, and chain sampling bias may also occur because heavy users may be more prone to share information about the study with other contacts than light users [16].

Self-selection bias (individuals who select themselves for the survey) may be of great importance [12], notably, in consideration of the usually relatively low participation rate [16]. It is difficult to estimate the impact of any selection bias because information on non-participants is usually not available, and comparisons between the included and the excluded samples are not feasible [12].

There is some evidence that the self-selected samples of Internet-based surveys may systematically differ from samples drawn from the general population using other sampling procedures [16,17]. One study showed that an Internet-based study sample had higher past month rates of alcohol and marijuana use than those found in other similar and non-self-selected samples of smokers (behaviors also possibly more easy to disclose online) [17]. Similarly, comparison of registry-recruited cancer survivors with an online recruited sample found that the Internet sample has lower social support and greater mood disturbances than the cancer-based registry-recruited one [18]. Another study found also that participants who preferred online surveys to paper-pencil questionnaires differed from their counterparts on a number of sociodemographic variables [19]. The effect of the self-selection bias is also possibly important in large sample size studies, as

suggested from differences in actual election results and a number of online opinion surveys [12].

To the best of our knowledge, studies are lacking on the possible differences between "pure or perfect" random samples and self-selected samples of users of specific Internet services such as online games or social network websites. This weakness is possibly explained by the difficulty for researchers, independently from websites owners, to obtain random samples of Internet users on specific websites.

The online game World of Warcraft (WoW) offers some possibilities to approach the question of selection and self-selection bias in online surveys. WoW gamers have been specifically studied in online self-selected surveys in attempts to assess motivations to play and possible psychological factors associated with gaming addiction [20-23].

In WoW, players assume the role of a fictional character, or "avatar". An avatar is characterized by a number of elements such as name and visual representation. The avatar's progression is a core attribute of WoW, implying that an avatar will develop new skills and powers as rewards for the success obtained during in-game missions or quests (eg, beating a monster, finding something specific, exploring areas of the game). Each avatar's progression is accessible via the "Armory", an official database reporting the achievements related to each avatar evolving in WoW [24]. Players commonly regroup themselves in guilds (hierarchical organizations of avatars with shared objectives and backgrounds). Each guild has its particular regulations. Players who want to join a given guild usually need to contact the guild's chief and explain their motivations to join the guild and to give some evidence that their avatar meets the guild's conditions [22].

Furthermore, the psychological characteristics of the gamers, such as motivation to play [21,23,25], have been shown to be associated with actual in-game behaviors and achievements as reported by the Armory scores [22]. Accordingly, the Armory scores of a given avatar, to some extent, reflect the game style and commitment of a given WoW player (ie, details of the achievements reached). The variables assessed in the present study were extracted from the armory and could be considered as an "ecological measure" (the measures automatically collected during game play represent direct in-game behaviors) of both the commitment of the players and their playing preferences [22].

WoW thus offers the possibility of comparing characteristics of the progression of self-selected avatars to a "pure or perfect" random sample of avatars. A sample is considered pure or perfect in the sense that every randomly selected avatar is actually included in the sample, whereas in classic studies (ie, non-self-selected samples), subjects are allowed to refuse to participate, which could induce a selection bias.

The aim of the current study was to compare the armory characteristics of two self-selected samples with a random sample of WoW avatars.

Methods

Summary

The study compared a random sample of avatars with two different self-selected samples. Only the avatars at the maximum level of the game version at inclusion were included in the study. The mechanics of "leveling" is as follows: When a player starts to play with an avatar, this avatar automatically starts at level 1. While playing, avatars gain experience points and these points allow the avatar to reach new levels (10,000 points to reach level 2; 25,000 to reach level 3, etc). Each new game version allows avatars to gain higher maximum levels. At the creation of the game, the maximum level an avatar could reach was 60. Each time that an expansion pack is released, the highest reachable level is raised (80 for Wrath of the Lich King and 85 for the Cataclysm versions of the game).

As described below, most of the avatars of the self-selected samples were at the maximum level. This level is not the maximum of the possible in-game achievements (reflected by the armory scores) but something like a mandatory "pass" for certain important tasks in the game (especially raids). To be considered as a seriously involved player, an avatar has to reach this maximum. So, including only avatars at this maximum level makes the avatars more comparable in terms of game "involvement".

First Self-Selected Sample

The first self-selected sample (self-selected sample 1) of avatars was from a study on the relationships between players' self-reported motives to play and their in-game behaviors [21]. The study was performed between June and December 2010 and was approved by the ethical committee of the Psychology Department of the University of Geneva. Inclusion criteria were French-speaking WoW players who were aged 18 years or older. Participants were recruited through advertisements posted in dedicated French-language forums: a guilds forum, an official Blizzard WoW forum, and more general online and video games forums. Some participants also joined the study after having heard about it in the local press or from television interviews. All participants gave online consent prior to starting the online survey. So, the sample included the avatars of online gamers who actively participated in the study given the identity of their avatars. Concomitant avatars' in-game behaviors were collected through the French Armory website [24].

The WoW avatar achievements studied here and reported in the Armory (Figure 1) are as follows: general achievements, quests (progression in the various available quests in the game), exploration (exploring each area in the game), player versus player (fighting other players), dungeons and raids (raids or dungeon crawling, ie, specific missions needing a group of players to achieve a common objective), profession, reputation, word events, and total completed. These achievement scores (Figure 1) were reported in the following format: score, maximum score (maximum possible score), and percentage of the maximum score. This percentage is calculated as the ratio of the scores gained by the player to the maximum possible score of the achievement in question multiplied by 100. For instance, a player who has obtained a score of 20 for quests out of a maximum of 127 is credited with a percentage of 15.75. All other percentages considered for the analysis were also calculated in the same way, and their means were compared across samples. Taking the percentage allows comparison of the avatars in different time periods despite possible modification in the WoW game. Some achievements such as "feats of strengths" and "total points" were not expressed as a percentage and were therefore not included in the study because of the difficulty in interpreting these scores in the case of game modifications.

Of a total of 1601 participants who started the survey (self-selection), 690 completed it (43.10%) and concomitantly provided the names of their avatar and the realms in which they play (ie, the name of their server, which is necessary to identify the avatar). Among these avatars, only those with a level of 80 were included in the present study, which represents 663 avatars of 690 (96.1%). This level was the higher one at the time of inclusion of the self-selected sample.

Figure 1. Example from the Armory.
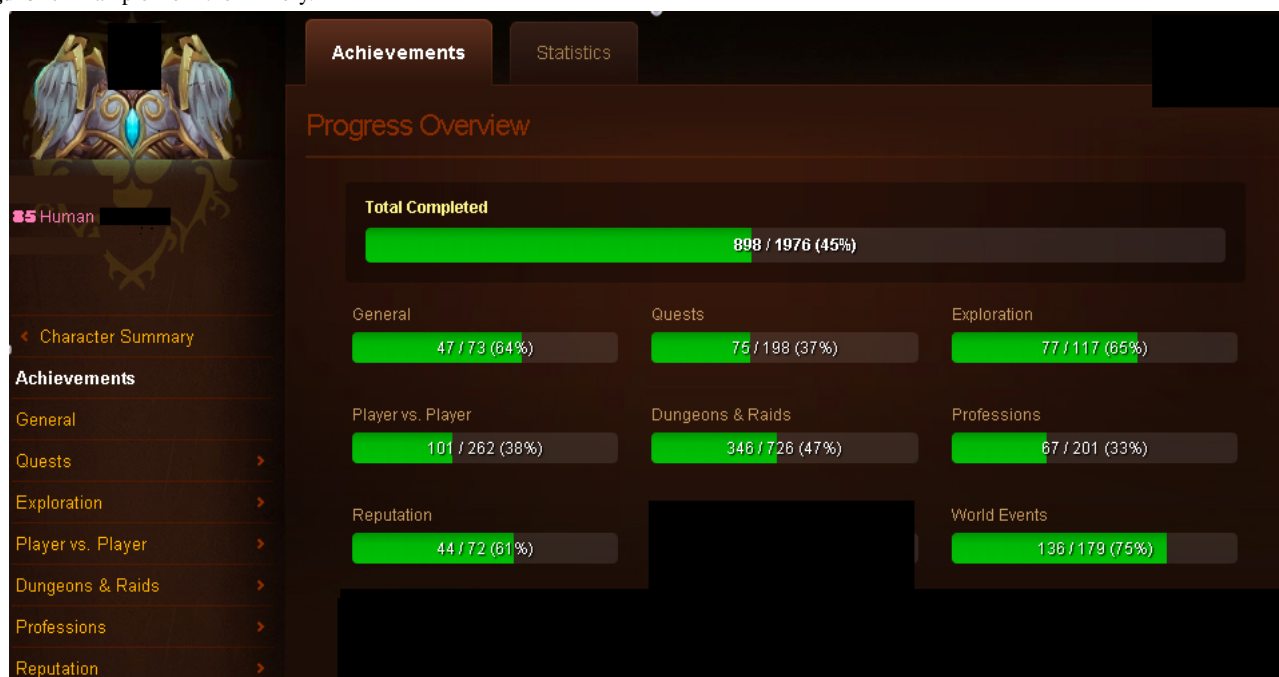
Second Self-Selected Sample

The second self-selected sample (self-selected sample 2) of avatars was from a study performed between December 2011 and April 2012. Similarly, the sample included the avatars of online gamers who actively participated in the study given the identity of their avatars. Concomitant avatars' in-game behaviors were collected in the same way as for the first self-selected sample.

The study, approved by the ethical committee of the Psychology Department of the University of Geneva, had similar purposes to the first study and similar recruitment procedures. The sample was assessed with the same measures from the Armory. One important added value of this smaller sample was the time of recruitment, which took place after the release of the Cataclysm version of the game in December 2010; it is therefore a version of the game similar to the one related to the random sample described in the next section. At that time, the maximum level was raised from 80 to 85.

Furthermore, response bias is considered as an "individual" characteristic of a given sample; replication of the results on different samples could be considered as a way to increase the validity of a given study conclusion [26].

In total, 104 participants participated in the survey; of these, 99 avatars (95.2%) had a level of 85 (the maximum) and were included in the present study.

Random Sample of Avatars

The list of avatars was found on a public website of game players [27] on February 25, 2012. On this website [27], each server was presented with numbered avatars.

Due to the fact that the two self-selected samples were recruited among French speaking WoW players, only the French-speaking population of avatars (found on French servers) was considered in order to ensure group comparability.

Given that most of the avatars of the two self-selected samples (96.1% of the first one, 637/663 avatars, and 95.2% of the second one, 94/99 avatars) were at the maximum level of the game version at the time of recruitment, only those avatars were included in our study and compared with a random sample of avatars. For the sake of comparability, all avatars included in the study were at the 85 level, which is the maximum level related to the Cataclysm version.

To form the random sample, 600 avatars were randomly drawn. The number of selected avatars from each server was proportional to the contribution of each server to the total population. The random allocation was made using a specialized website [28]. The avatars selected by this procedure were then searched for and assessed in the WoW official comprehensive database [24]. Only avatars considered as still active by the WoW game were registered in this database.

Statistical Analyses

Statistical analyses were performed using SPSS, version 18.0. An initial exploratory analysis involved the calculation of percentages, as well as means and standard deviations of the above-mentioned outcome measures.

To address the research question, analysis of variance (ANOVA) or t tests are appropriate in comparing mean percentages across groups. However, although F tests are robust against departure from normality, the homogeneity of variance is a strong assumption that must be satisfied for the ANOVA results to be reliable. As percentages and proportions variables are not likely to meet normality and homogeneity of variance assumptions, the arcsine-transformation is often used with this type of data and serves the purpose of normalizing them and stabilizing their variance. First, all variables were expressed as proportions, that is, between 0 and 1, and then they were transformed according to the following formula: $y=2\arcsin(\mathrm{SQRT}(p))$, where p stands for proportion. The combined effect of the square root with the

inverse sine compresses the upper tail of the distribution and stretches out both tails relative to the middle. The ANOVAs and *t* tests were done on the transformed variables after visual inspections of their normality and Q-Q plots excluded inacceptable patterns. But for the sake of completeness, the tables also display the variables in their original scale. In a first step, the three samples were compared (self-selected sample 1 vs self-selected sample 2 vs random sample) using a one-way between-group ANOVA to explore the impact of each sample on a list of 10 selected variables. In a second step, the two self-selected samples were merged, since they did not differ (shown by post hoc comparison tests), and two-sample *t* tests were done, comparing the random sample with the merged self-selected samples. To account for multiple comparisons testing (10 multiple ANOVAs and *t* tests), we performed Bonferroni corrections deflating alpha type I error so that the adjusted significance level is alpha/10 (here .005). Finally, a

chi-square test was carried out to compare proportions of avatar per guild between the two types of samples.

Results

Table 1 shows the ANOVA results for each variable of interest. Except for "guild membership", overall, statistically significant differences at the .005 level between the three samples were found for all assessed variables. Bonferroni post hoc tests showed that these differences mainly occurred within each self-selected sample compared with the random sample. However, these differences were not statistically significant between the self-selected sample 1 and the random sample for exploration (*P*=.6), between the self-selected sample 2 and the random sample for total completed (*P*=.1), and between neither one of the self-selected samples and the random sample for guild (*P*=1.0 respectively).

Table 1. Comparison of mean values between three samples: self-selected sample 1 (n=663), self-selected sample 2 (n=99), and random sample (n=478) by one-way ANOVA performed on the transformed variables.

WoW characteristics	Variables successively presented in their original and transformed values	Self-selected sample 1, mean (SD)	Self-selected sample 2, mean (SD)	Random sample, mean (SD)	*P* value
Dungeons and raids	Original mean proportion	0.40 (0.20)	0.42 (0.24)	0.28 (0.18)	
	Arcsine-transformed value	1.36 (0.44)	1.38 (0.54)	1.07 (0.43)	<.001
Word events	Original mean proportion	0.40 (0.33)	0.43 (0.34)	0.21 (0.25)	
	Arcsine-transformed value	1.34 (0.81)	1.38 (0.83)	0.85 (0.61)	<.001
Exploration	Original mean proportion	0.62 (0.35)	0.76 (0.26)	0.64 (0.27)	
	Arcsine-transformed value	1.91 (0.85)	2.23 (0.66)	1.91 (0.67)	<.001
General	Original mean proportion	0.60 (0.18)	0.60 (0.18)	0.48 (0.15)	
	Arcsine-transformed value	1.81 (0.44)	1.81 (0.44)	1.54 (0.32)	<.001
Player versus player	Original mean proportion	0.35 (0.19)	0.27 (0.18)	0.20 (0.14)	
	Arcsine-transformed value	1.24 (0.42)	1.04 (0.44)	0.88 (0.37)	<.001
Profession	Original mean proportion	0.48 (0.29)	0.42 (0.28)	0.27 (0.23)	
	Arcsine-transformed value	1.54 (0.67)	1.40 (0.68)	1.02 (0.56)	<.001
Quests	Original mean proportion	0.50 (0.25)	0.40 (0.27)	0.26 (0.19)	
	Arcsine-transformed value	1.61 (0.61)	1.38 (0.69)	1.04 (0.43)	<.001
Reputation	Original mean proportion	0.35 (0.25)	0.39 (0.30)	0.23 (0.22)	
	Arcsine-transformed value	1.24 (0.60)	1.33 (0.75)	0.94 (0.57)	<.001
Total completed	Original mean proportion	0.43 (0.19)	0.54 (0.27)	0.37 (0.21)	
	Arcsine-transformed value	1.42 (0.42)	1.73 (0.70)	1.31 (0.50)	<.001
Involvement in guilds	Original mean proportion	0.87 (0.34)	0.86 (0.35)	0.88 (0.32)	
	Arcsine-transformed value	2.73 (1.06)	2.70 (1.10)	2.77 (1.01)	.7

Comparing the two self-selected samples, Bonferroni's post hoc tests showed that they differed on the following variables: exploration, player versus player, and quests (*P*<.001 respectively). No further difference was observed for the other variables. We merged these two self-selected samples into one bigger sample, which in turn was compared to the random

sample. Table 2 shows the two-sample *t* tests between the new self-selected sample and the random sample. Both samples differed significantly at the .005 level on each variable, except for guild membership and exploration, with the random sample having a similar mean to that of the self-selected sample.

Table 2. Comparison of mean values between the merged self-selected sample (n=762) and the random sample (n=478) by *t* test performed on the transformed variables.

WoW characteristics	Variables are successively presented in their original and transformed values	All self-selected sample players, mean (SD)	Random sample, mean (SD)	Mean difference (99.5% CI)	*P* value
Dungeons and raids	Original mean proportion	0.40 (0.20)	0.28 (0.18)		
	Arcsine-transformed value	1.36 (0.45)	1.07 (0.43)	0.29 (0.22-0.36)	<.001
Word events	Original mean proportion	0.40 (0.33)	0.21 (0.25)		
	Arcsine-transformed value	1.34 (0.81)	0.85 (0.61)	0.49 (0.38-0.60)	<.001
Exploration	Original mean proportion	0.64 (0.34)	0.64 (0.27)		
	Arcsine-transformed value	1.95 (0.83)	1.91 (0.67)	0.05 (-0.08 to 0.17)	.3
General	Original mean proportion	0.60 (0.18)	0.48 (0.15)		
	Arcsine-transformed value	1.81 (0.44)	1.54 (0.32)	0.27 (0.21-0.33)	<.001
Player versus player	Original mean proportion	0.34 (0.19)	0.20 (0.14)		
	Arcsine-transformed value	1.21 (0.43)	0.88 (0.37)	0.33 (0.26-0.39)	<.001
Profession	Original mean proportion	0.48 (0.29)	0.27 (0.23)		
	Arcsine-transformed value	1.52 (0.67)	1.02 (0.56)	0.50 (0.40-0.60)	<.001
Quests	Original mean proportion	0.49 (0.25)	0.26 (0.19)		
	Arcsine-transformed value	1.58 (0.62)	1.04 (0.43)	0.55 (0.46-0.63)	<.001
Reputation	Original mean proportion	0.36 (0.26)	0.23 (0.22)		
	Arcsine-transformed value	1.26 (0.63)	0.93 (0.57)	0.32 (0.23-0.42)	<.001
Total completed	Original mean proportion	0.44 (0.21)	0.37 (0.21)		
	Arcsine-transformed value	1.46 (0.48	1.31 (0.50)	0.15 (0.07-0.22)	<.001
Involvement in guilds	Original mean proportion	0.87 (0.34)	0.88 (0.32)		
	Arcsine-transformed value	2.73 (1.06)	2.77 (1.01)	-0.04 (-0.21 to 0.12)	.5

Some guilds were overrepresented in the self-selected samples (a form of guild effect: people from the same guild may encourage their partners to participate in the study). Table 3 shows the number of avatars per guild. The sample sizes here are lower since not every avatar belongs to a guild. The range is between 1 and 11, with one guild from the self-selected sample having 11 participating avatars. A chi-square test reveals that the distribution of avatar per guild is different between the two types of samples.

Table 3. Observed (expected) number of avatars per guild: random sample (n=421) and self-selected samples (n=662).

Number of avatars per guild	Random sample[a]	All self-selected sample	χ^2_2	*P* value
1	363 (345.0)	393 (411.0)		
2	29[b] (35.6)	49 (42.4)		
3-11	0 (11.9)	26 (14.1)	25.8	<.001

[a]Number of avatars with a guild affiliation.

[b]29 guilds with 2 avatars per guild.

Discussion

Principal Findings

In the French-speaking community of WoW players, three samples of avatars, one purely random and two self-selected, were used to assess the potential self-selection bias of Internet-based studies. To our knowledge, this is the first study to include a perfect random sample, since all randomly selected subjects (avatars) were incorporated in the sample.

The method used in this paper is somewhat new, dealing with the opportunity given by the development of online avatars of Internet users. Table 4 gives some details about the similarities and the differences related to Internet surveys, surveys on online gamers, and studies on avatars.

Table 4. Comparison of Internet surveys, online game surveys, and studies on avatars.

	Internet survey	Survey on online gamers	Study on avatars from a self-selected sample	Study on avatars from a random sample
Included	Internet users	On-line game users	Fictive character (video game players' avatar) linked to a given user who self-selected for study participation (as in the two self-selected samples of our study)	Video game players' avatar selected randomly from a database including all characters (like the random sample of our study)
Active participation in the survey	Mandatory for study participation	Mandatory for study participation	The characteristics of a given avatar are drawn from the Web. Active participation is linked to the fact that the user decided to disclose their avatar's information for the study.	No possible active participation to a survey. The characteristics of a given avatar are drawn from the Web (like the achievements characteristics of the random sample).
Possible data obtained	Data on the participant (human) Internet user profile (eg, psychological measures, reported Internet use)	Data on the participant (human) Internet-game user (eg, psychological measures, reported Internet use)	Data automatically collected related to the avatars (eg, achievements) or data chosen by the player (eg, avatar name, gender, guild affiliation) like in our study for the self-selected samples. It remains possible to link the characteristics of the avatar with the user who self-selected him or herself (not available in our study).	Data automatically collected related to the avatars (eg, achievements) or data chosen by the player (eg guild affiliation) like in our study for the random sample.
Self-selection bias	Possible self-selection bias. The participant (human) is or is not informed about the survey and decides to or not to participate and to complete the survey.	Possible self-selection bias. The participant (human) is or is not informed about the survey and decides to or not to participate and to complete the survey.	The human user decided to include their avatar in a study (self-selection bias, like for the 2 self-selected samples of our study)	The avatar is selected randomly from a database (like for the random sample of our study).

The samples were compared on the basis of the in-game achievements of the avatars expressed in percentages. This allows comparison of avatars despite possible game modifications, as shown by the lack of differences between the two self-selected samples recruited at two different times. The second self-selected sample and the random sample were both included during the same Cataclysm version of the game, whereas the first self-selected sample was included before this game version.

According to the hypothesis of a self-selection bias, it appears from the study results that a self-selected sample of website users differs from a "pure or perfect" random sample. The self-selected samples had higher scores than the random sample on most of the assessed in-game behavior variables. The self-selected samples appear to be more involved in the game than the random sample avatars. This could occur for different reasons.

To self-select, a player needs first to see the advertisement for a study (eg, "We are looking for active World of Warcraft players, >18 years old, to participate in an online survey on your motives to play and your psychological profile. We will ask the name of your main avatar and match your answers to Blizzard armory's data. The questionnaire will take approximately 15 minutes of your time"). Second, a player needs to consider participating and to agree to it. Therefore, having the will to be involved in a study could lead to the selection of specific subjects with certain characteristics (eg, personality, game involvement, special interest in the purpose of the study). Because the participants responded to an Internet advertisement for the study, highly involved players are more likely to see the ad than occasional players because of the time spent on WoW-related websites. On the other hand, one could also assume that the will to participate is related to both involvement in the game and an interest in the proposed studies.

The study finding is consistent with those of other studies linking survey participation with involvement (greater interest, connection, and concern related to the given behavior or possibly to the study results) in the assessed behaviors [19,29].

One may hypothesize that the statistically significant results of the study were due to the large sample size and that type I error (finding a difference when in fact there is none) could not be ruled out. But the magnitude of the differences found between the groups (as displayed in Table 2) cannot be imputed to chance alone and therefore does not support this point of view.

Limitations

Some limitations warrant further consideration. First, the specificity of WoW, including the guild effect (players organized in a guild), may increase the inclusion of participants who are highly involved in the game via a chain sampling bias. This may limit the generalizability of the results to other domains of Internet use–related behaviors. However, most Internet-related activities involve some form of social networking that promotes chain sampling activities and a possibly similar bias.

Second, the study was done on avatars and not directly on people. People may have more than one avatar on WoW. Thus, we cannot exclude the possibility that the results may be partly explained by differences between the avatar chosen by the self-selected sample as representative and the randomly selected avatars. Inclusion of active avatars at the maximum level related to each game extension and particularly lack of statistically significant differences in guild involvement and exploration (for the first self-selected sample) and on total completed (for the second self-selected sample) suggest, however, that the random sample is composed of at least reasonably credible avatars involved in the game.

Although most of the avatars were affiliated with guilds, affiliation was not an inclusion criterion. However, it was used to assess a form of guild effect (ie, higher proportions of avatars from the same guilds in the self-selected groups in comparison to the random one). Furthermore, guild participation could be considered as a useful index of "serious" avatar in-game activity (ie, the avatar was accepted by a guild).

Conclusions

Because of the important differences between the self-selected samples and the randomly selected sample, and despite the acknowledged limitations, the study invites careful consideration of the conclusions made from online self-selected samples and the possibility of an overrepresentation of subgroups of more involved or more concerned users.

Therefore, it does not appear possible to draw general epidemiological conclusions from Internet-based self-selection surveys (eg, on the prevalence of game addiction among website users or the general population). However, the studies may be of high interest to subgroups of users who are more involved in the game and the study purpose. In particular, such studies may allow the linking together of different assessed variables (such as mood, motives, or personality and a given behavior) in the studied sample. This remains important, particularly because of the possible advantages of online studies (eg, large sample sizes, possible access to people who are usually more difficult to reach, access to stigmatized behaviors).

The possible collaboration with webmasters may further improve understanding of the representativeness of self-selected samples by the random selection of the users (ie, contacting users by email to build a random sample as control group) or by comparison of the responders to non-responders regarding general characteristics such as features related to website use or, to some extent, potential biases regarding clinical variables (eg, game addiction).

Conflicts of Interest

None declared.

References

1. Schonlau M. Will web surveys ever become part of mainstream research? J Med Internet Res 2004 Sep 23;6(3):e31 [FREE Full text] [doi: 10.2196/jmir.6.3.e31] [Medline: 15471757]

2. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. J Epidemiol Community Health 2003 Jan;57(1):68-73 [FREE Full text] [Medline: 12490652]

3. McCabe SE, Boyd CJ, Couper MP, Crawford S, D'Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. J Stud Alcohol 2002 Nov;63(6):755-761. [Medline: 12529076]

4. McCabe SE, Boyd CJ, Young A, Crawford S. Feasibility study for collecting alcohol and other drug use data among secondary school students: a web-based survey approach. J Drug Educ 2004;34(4):373-383. [Medline: 16117249]

5. McCabe SE. Comparison of web and mail surveys in collecting illicit drug use data: a randomized experiment. J Drug Educ 2004;34(1):61-72. [Medline: 15468748]

6. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004;59(2):93-104. [doi: 10.1037/0003-066X.59.2.93] [Medline: 14992636]

7. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. J Med Internet Res 2011;13(4):e109 [FREE Full text] [doi: 10.2196/jmir.1593] [Medline: 22155721]

8. Ritter P, Lorig K, Laurent D, Matthews K. Internet versus mailed questionnaires: a randomized comparison. J Med Internet Res 2004 Sep 15;6(3):e29 [FREE Full text] [doi: 10.2196/jmir.6.3.e29] [Medline: 15471755]

XSL•FO

RenderX

9.   Kypri K, Gallagher SJ, Cashell-Smith ML. An internet-based survey method for college student drinking research. Drug Alcohol Depend 2004 Oct 5;76(1):45-53. [doi: 10.1016/j.drugalcdep.2004.04.001] [Medline: 15380288]

10. McCabe SE. Screening for drug abuse among medical and nonmedical users of prescription drugs in a probability sample of college students. Arch Pediatr Adolesc Med 2008 Mar;162(3):225-231 [FREE Full text] [doi: 10.1001/archpediatrics.2007.41] [Medline: 18316659]

11. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]

12. Bethlehem J. Selection bias in web surveys. International Statistical Review 2010;78(2):161-188. [doi: 10.1111/j.1751-5823.2010.00112.x]

13. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775 [FREE Full text] [doi: 10.1093/ntr/ntq086] [Medline: 20530194]

14. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]

15. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009(3):MR000008. [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]

16. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]

17. Ramo DE, Prochaska JJ. Prevalence and co-use of marijuana among young adult cigarette smokers: An anonymous online national survey. Addict Sci Clin Pract 2012;7(1):5 [FREE Full text] [doi: 10.1186/1940-0640-7-5] [Medline: 23186143]

18. Owen JE, Bantum EO, Criswell K, Bazzo J, Gorlick A, Stanton AL. Representativeness of two sampling procedures for an internet intervention targeting cancer-related distress: a comparison of convenience and registry samples. J Behav Med 2013 May 4:-. [doi: 10.1007/s10865-013-9509-6] [Medline: 23645145]

19. Mayr A, Gefeller O, Prokosch HU, Pirkl A, Fröhlich A, de Zwaan M. Web-based data collection yielded an additional response bias--but had no direct effect on outcome scales. J Clin Epidemiol 2012 Sep;65(9):970-977. [doi: 10.1016/j.jclinepi.2012.03.005] [Medline: 22742911]

20. Zanetta Dauriat F, Zermatten A, Billieux J, Thorens G, Bondolfi G, Zullino D, et al. Motivations to play specifically predict excessive involvement in massively multiplayer online role-playing games: evidence from an online survey. Eur Addict Res 2011;17(4):185-189. [doi: 10.1159/000326070] [Medline: 21494046]

21. Kuss DJ, Louws J, Wiers RW. Online gaming addiction? Motives predict addictive play behavior in massively multiplayer online role-playing games. Cyberpsychol Behav Soc Netw 2012 Sep;15(9):480-485. [doi: 10.1089/cyber.2012.0034] [Medline: 22974351]

22. Billieux J, Van der Linden M, Achab S, Khazaal Y, Paraskevopoulos L, Zullino D. Why do you play World of Warcraft? An in-depth exploration of self-reported motivations to play online and in-game behaviours in the virtual world of Azeroth. Comput Human Behav 2013;29(1):103-109.

23. Yee N. Motivations for play in online games. Cyberpsychol Behav 2006 Dec;9(6):772-775. [doi: 10.1089/cpb.2006.9.772] [Medline: 17201605]

24. World of warcraft-Site officiel. URL: http://eu.battle.net/wow/fr/ [accessed 2014-06-26] [WebCite Cache ID 6Qd2EhVIB]

25. Billieux J, Chanal J, Khazaal Y, Rochat L, Gay P, Zullino D, et al. Psychological predictors of problematic involvement in massively multiplayer online role-playing games: illustration in a sample of male cybercafé players. Psychopathology 2011;44(3):165-171. [doi: 10.1159/000322525] [Medline: 21372629]

26. Jonathon R, Halbesleben B, Whitman M. Evaluating Survey Quality in Health Services Research: A Decision Framework for Assessing Nonresponse Bias. Health Serv Res 2013 Jun;48(3):913-930. [doi: 10.1111/1475-6773.12002]

27. WoW heroes. URL: http://www.wow-heroes.com/ [accessed 2014-06-19] [WebCite Cache ID 6QRsUT2cG]

28. Research randomizer. URL: http://www.randomizer.org/ [accessed 2014-06-19] [WebCite Cache ID 6QRsWJ2Bq]

29. Cranford JA, McCabe SE, Boyd CJ, Slayden J, Reed MB, Ketchie JM, et al. Reasons for nonresponse in a web-based survey of alcohol involvement among first-year college students. Addict Behav 2008 Jan;33(1):206-210 [FREE Full text] [doi: 10.1016/j.addbeh.2007.07.008] [Medline: 17728069]

Abbreviations

ANOVA: analysis of variance
CI: confidence interval
WoW: World of Warcraft

XSL•FO
RenderX

XSL•FO
RenderX

Original Paper

Longitudinal Accuracy of Web-Based Self-Reported Weights: Results From the Hopkins POWER Trial

Gerald J Jerome[1,2], PhD; Arlene Dalcin[2], RD; Janelle W Coughlin[3], PhD; Stephanie Fitzpatrick[4], PhD; Nae-Yuh Wang[2], PhD; Nowella Durkin[2]; Hsin-Chieh Yeh[2], PhD; Jeanne Charleston[2], RN; Thomas Pozefsky[5], MD; Gail L Daumit[2], MD, MHS; Jeanne M Clark[2], MD, MPH; Thomas A Louis[6], PhD; Lawrence J Appel[2], MD, MPH

[1]Department of Kinesiology, Towson University, Towson, MD, United States

[2]Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[3]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[4]Department of Preventive Medicine, Rush University Medical Center, Chicago, IL, United States

[5]Division of Endocrinology and Metabolism, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[6]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Gerald J Jerome, PhD
Department of Kinesiology
Towson University
8000 York Road
Towson, MD, 21252
United States
Phone: 1 410 704 5283
Fax: 1 410 704 3912
Email: gjerome@towson.edu

Abstract

Background: Websites and phone apps are increasingly used to track weights during weight loss interventions, yet the longitudinal accuracy of these self-reported weights is uncertain.

Objective: Our goal was to compare the longitudinal accuracy of self-reported weights entered online during the course of a randomized weight loss trial to measurements taken in the clinic. We aimed to determine if accuracy of self-reported weight is associated with weight loss and to determine the extent of misclassification in achieving 5% weight loss when using self-reported compared to clinic weights.

Methods: This study examined the accuracy of self-reported weights recorded online among intervention participants in the Hopkins Practice-Based Opportunities for Weight Reduction (POWER) trial, a randomized trial examining the effectiveness of two lifestyle-based weight loss interventions compared to a control group among obese adult patients with at least one cardiovascular risk factor. One treatment group was offered telephonic coaching and the other group was offered in-person individual coaching and group sessions. All intervention participants (n=277) received a digital scale and were asked to track their weight weekly on a study website. Research staff used a standard protocol to measure weight in the clinic. Differences (self-reported weight – clinic weight) indicate if self-report under (-) or over (+) estimated clinic weight using the self-reported weight that was closest in time to the clinic weight and was within a window ranging from the day of the clinic visit to 7 days before the 6-month (n=225) and 24-month (n=191) clinic visits. The absolute value of the differences (absolute difference) describes the overall accuracy.

Results: Underestimation of self-reported weights increased significantly from 6 months (mean -0.5kg, SD 1.0kg) to 24 months (mean -1.1kg, SD 2.0kg; *P*=.002). The average absolute difference also increased from 6 months (mean 0.7kg, SD 0.8kg) to 24 months (mean 1.3, SD 1.8kg; *P*<.001). Participants who achieved the study weight loss goal at 24 months (based on clinic weights) had lower absolute differences (*P*=.01) compared to those who did not meet this goal. At 24 months, there was 9% misclassification of weight loss goal success when using self-reported weight compared to clinic weight as an outcome. At 24 months, those with self-reported weights (n=191) had three times the weight loss compared to those (n=73) without self-reported weights (*P*<.001).

Conclusions: Underestimation of weight increased over time and was associated with less weight loss. In addition to intervention adherence, weight loss programs should emphasize accuracy in self-reporting.

XSL•FO

RenderX

KEYWORDS

self-report; weight loss; obesity; Internet

Introduction

The use of technology to support lifestyle-based treatment of obesity is commonplace in both research studies and commercial programs [1,2]. Specifically, self-monitoring of weight is a frequently recommended weight-loss strategy that can be facilitated by either websites or mobile phone apps [3]. Although self-monitoring of weight is a commonly identified, evidence-based strategy for weight loss, both technology-based and traditional paper-based cross-sectional studies indicate that individuals underestimate their weight and that the accuracy can vary for different demographic groups [4-8].

Few studies have examined the accuracy of self-reported weight over time. A 12-week weight management study (N=27) found self-reported weights recorded on a mobile phone were underestimated but were strongly correlated with weights taken in the clinic at both baseline and 12 weeks [9]. Results from a 6-month weight-loss trial (N=234) indicated that self-reported weights underestimated observed weights and those participants who lost more weight had more accurate self-report [10]. The accuracy of self-reported weight has not been examined in the context of a weight-loss study beyond 6 months.

The current study compared self-reported weight from a study website to clinic weights at 6-month and 24-month follow-up in the Hopkins Practice-Based Opportunities for Weight Reduction (POWER) trial, a three-arm randomized weight-loss trial with gender and race diversity [11,12]. The study also examined the association of accuracy of self-reported weight with the extent of weight loss and determined the rate of misclassification in achieving 5% weight loss when using self-reported compared to clinic weights as the follow-up weight.

Methods

Overview

The POWER trial at Hopkins was a randomized trial examining the effectiveness of two lifestyle-based weight-loss interventions (n=277) compared to a control group (n=138) among obese adult patients at six primary care practices [11,12]. Participants were at least 22 years of age, with a body mass index (BMI) ≥30 kg/m^2 and at least one additional cardiovascular risk factor. An institutional review board approved the study and all participants provided informed consent. Detailed descriptions of intervention design, methods, and main results have been published [11,12].

Intervention Summary

Intervention participants had a 5% weight-loss goal and access to a study website that included learning modules and tools for self-monitoring weight, caloric intake, and exercise. They also received a digital scale for home use and directions to (1) weigh themselves at the same time of day, in the same amount of clothing with the same scale, and (2) enter this weight on the study website at least weekly while trying to lose weight and daily during weight maintenance.

During the first 6 months, those in the Call-Center Directed intervention were offered 15 coaching calls and those in the In-Person Directed intervention were offered 21 groups and nine individual coaching sessions. From Month 7 to the end of the study, call-center participants were offered monthly calls and in-person participants were offered both individual and group sessions monthly.

Primary Outcome

As previously reported, mean change in clinic measured weight (24 months – baseline) was –0.8 kg in the control arm, –4.6 kg in the call-center arm (*P*<.001 compared to control), and –5.1 kg in the in-person arm (*P*<.001 compared to control) [11]. There was no significant difference in weight loss between call-center directed and in-person directed arms. At 24 months, 37.9% (105/277) of the intervention participants achieved 5% weight loss.

Measures

Demographics were self-reported at baseline. The number of completed coaching contacts and weight log-ins were recorded. Trained research staff, masked to intervention assignment, measured weights in the clinic at the randomization visit (baseline) and at the 6-month and 24-month follow-up visits. Following a standard protocol that included removing shoes and emptying pockets, weight was measured on a high-quality calibrated digital scale with the participant wearing light indoor clothes. Two weights were taken at each time point, and if needed, a third weight was taken to resolve any discrepancies. The assessment of weight during a clinic visit was independent of all intervention efforts. The accuracy of the clinic scales was verified annually by a third party.

Self-reported weight was based on the self-reported weight entered on the study website that was closest in time to the clinic weight and was within a window ranging from the day of the clinic visit to 7 days before the clinic visit. The intervention goal was to record a self-reported weight at least weekly, although the website allowed for daily self-report.

Analytic Plan

Weights are reported in kilograms. Differences (self-reported weight – clinic weight) indicate under (–) or over (+) estimation of clinic weight. The absolute value of the difference (absolute difference) describes the overall accuracy. We regressed absolute difference on weight change and controlled for age, sex, race, baseline BMI, intervention arm, number of coaching contacts

completed, number of weeks with a self-reported weight, and the difference in days between the self-reported and clinic weights.

In the main results paper, follow-up clinic weight was used to determine weight outcomes—both the amount of weight change and classifications of participants as either achieving or not achieving the study goal of 5% weight loss. In the current analyses, the same approach was used (ie, using clinic weights) to determine the gold standard for weight loss and subsequent classifications. We also used self-reported weights at 24-month follow-up to calculate an alternative weight loss and weight-loss classification. Misclassification refers to the alternative method (using self-reported follow-up) providing a different weight-loss classification compared to the gold standard (using clinic follow-up weight). We reported Cohen's kappa comparing classifications using clinic weight compared to self-reported weight at follow-up to determine weight change. A Bland-Altman plot is presented comparing clinic and self-reported weights at 24 months. The intraclass correlation between self-reported and clinic weight at 24 months was also reported.

Results

There were 277 participants in the active interventions. At 6 months, 28 did not have a self-reported weight within the target window and 24 were missing clinic weights resulting in 81.2% (225/277) with differences calculated. At 24 months, 83 did not have a self-reported weight within the target window and 13 were missing clinic weights resulting in 69.0% (191/277) with differences calculated.

Compared to those with missing data, the sample with differences calculated at 24 months had a higher percentage of men (*P*=.02), higher average contact completion (*P*<.001), and higher average number of self-reported weights (*P*<.001). Those with differences calculated at 24 months had an average age of 55 (SD 9.9) years, average BMI of 36 (SD 5.0) kg/m^2, average weight of 104 (SD 18.7) kg, 59.2% were women (113/191), and 38.7% were black (74/191).

The average number of contacts was 14 (SD 3.4) at 6 months and 30 (SD 9.3) at 24 months. The average number of weeks with at least one weight log-in was 22 (SD 5.12) at 6 months and 71 (SD 26.4) at 24 months.

At 6 months, 61.8% (139/225) of self-reported weights were lower, 22.2% (50/225) were equivalent, and 16.0% (36/225) were higher than the clinic weights. At 24 months, 77.5% (148/191) of self-reported weights were lower, 9.4% (18/191) were equivalent, and 13.1% (25/191) were above the clinic weights. As seen in Table 1, the degree of underestimation

increased from 6 months (mean −0.5kg, SD 1.0kg) to 24 months (mean −1.1kg, SD 2.0kg; *P*=.002). The average absolute difference also increased from 6 months (mean 0.7kg, SD 0.8kg) to 24 months (mean 1.3kg, SD 1.8kg; *P*<.001). Achieving the 5% weight loss goal was not associated with differences at 6 months (*P*=.24) or 24 months (*P*=.09). Participants who achieved 5% weight loss had lower absolute differences at 24 months (*P*=.01), but not at 6 months (*P*=.13) compared to those who did not meet this goal.

Figure 1 displays a distribution of the days between self-reported and clinic weights with a minimum score of −7 (self-reported weight was 7 days before the clinic weight) and a maximum score of 0 (self-reported and clinic weight were on the same day). At 6 months, 91.6% (206/225) of the self-reported weights were within 2 days of the clinic weight. At 24 months, 79.6% (152/191) of the self-reported weights were within 2 days of the clinic weight.

At 6 months, greater absolute difference between self-reported and clinic weight was associated with females, higher baseline BMI, less 6-month weight loss, fewer weeks with self-reported weights, and days between weights (Table 2). At 24 months, more weight loss and fewer days between weights were associated with smaller absolute differences. Significant associations among the independent and dependent variables were similar when examining differences rather than absolute differences as the outcome variable (data not shown).

When using self-reported weight at 24 months to calculate weight change from baseline clinic weight (see Table 3), there was 9% misclassification in achievement of weight-loss goal, 99% sensitivity, and 84% specificity compared to using clinic weights at 24 months. Cohen's kappa was used to determine level of nonrandom agreement comparing the use of self-reported weight to the use of clinic weight at 24 months to determine weight loss classifications (κ=.82). Those included in Table 3 had greater 24-month weight loss (mean −6.0kg, SD 9.2kg) compared to those who did not have a self-reported weight within the target window at 24 months (mean −2.1kg, SD 5.3kg; *P*<.001).

Figure 2 is a Bland Altman Plot comparing self-reported and clinic weights at 24 months. A majority of the extreme differences between self-reported and clinic weights occur between 80 kg and 130 kg. The intraclass correlation comparing self-reported and clinic weights at 24 months was .99.

Figure 3 displays differences between self-reported and clinic weights by the percent weight change at 24 months. Percent weight change was based on clinic weights. As seen in the figure, a majority of the self-reported weights were below compared to above the clinic weight. The spread of the differences increased as weight loss decreased.

Table 1. Difference in weights (self-reported – clinic) by attainment of 5% weight-loss goal at 6 and 24 months.

Difference	6 months			24 months		
	All (n=225)	Achieved 5% weight loss (n=112)	Did not achieve 5% weight loss (n=113)	All (n=191)	Achieved 5% weight loss (n=82)	Did not achieve 5% weight loss (n=109)
Difference[a] in kg, mean (SD)	−0.5 (1.0)	−0.4 (0.8)	−0.6 (1.1)	−1.1 (2.0)	−0.8 (1.2)	−1.2 (2.4)
Absolute difference[b] in kg, mean (SD)	0.7 (0.8)	0.6 (0.7)	0.8 (1.0)	1.3 (1.8)	1.0 (1.1)	1.5 (2.2)

[a]Difference (self-reported – clinic weight).

[b]Absolute difference = |self-reported weight – clinic weight|.

Table 2. Association of absolute difference in weights (self-reported – clinic) with weight change, demographic characteristics, and participation at 6 and 24 months.

Independent variable	Absolute difference[a]			
	6 months		24 months	
	B (SE)[b]	P	B (SE)[b]	P
Weight change[c] in kg	.04 (0.01)	<.001	.05 (0.02)	.01
Female	.38 (0.11)	.001	.19 (0.28)	.50
Black	.06 (0.11)	.58	.53 (0.28)	.06
Age	.003 (0.01)	.55	.02 (0.01)	.27
Body Mass Index (baseline)	.03 (0.01)	.02	.03 (0.03)	.31
Intervention arm	−.10 (0.11)	.36	−.45 (0.27)	.10
Coach contacts completed	−.00 (0.02)	.64	.02 (0.02)	.44
Self-reported weight, frequency	.04 (0.01)	.01	.00 (0.01)	.81
Days between weights	−.24 (0.04)	<.001	−.24 (.08)	.003

[a]Absolute difference = |self-reported weight – clinic weight|.

[b]B=beta coefficient, SE=standard error; adjusted for all variables listed.

[c]Weight change = follow-up clinic weight – baseline clinic weight.

Table 3. Weight loss goal attainment at 24 months based on final weight from clinic and self-reported weight (n=191).

Self-reported weight	Clinic weight	
	Achieved 5% weight loss, n (%)	Did not achieve 5% weight loss, n (%)
Achieved 5% weight loss	93 (48.7)	16 (8.4)
Did not achieve 5% weight loss	1 (0.5)	81 (42.4)

Figure 1. Time between self-reported and clinic weights at 6 and 24 months.
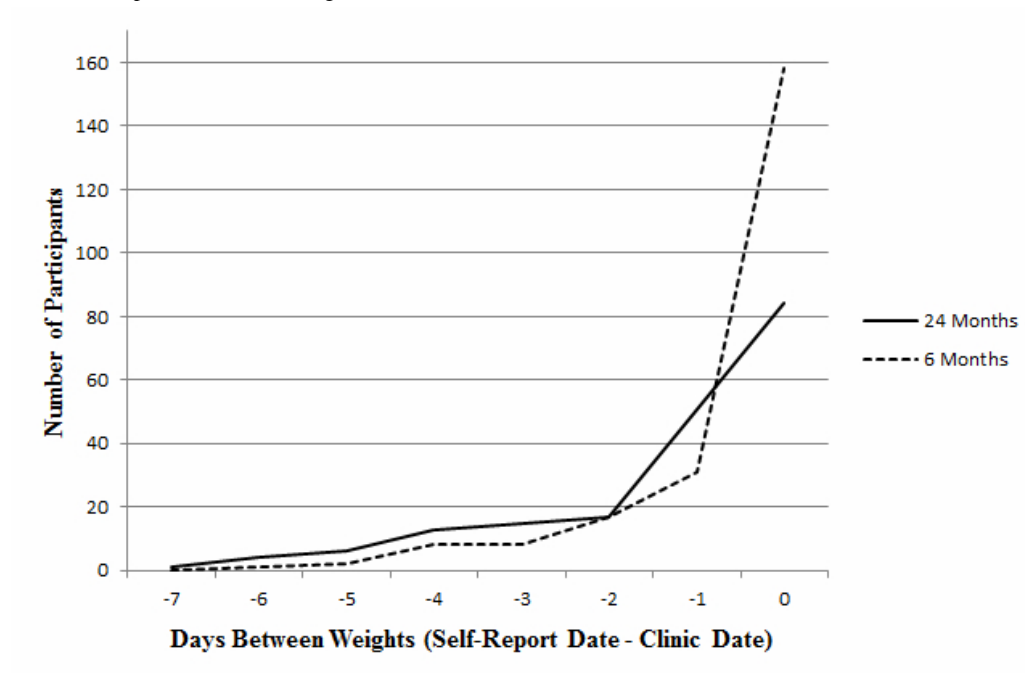
Figure 2. Differences in weights (self-reported – clinic) by the average of the two weights at 24 months.
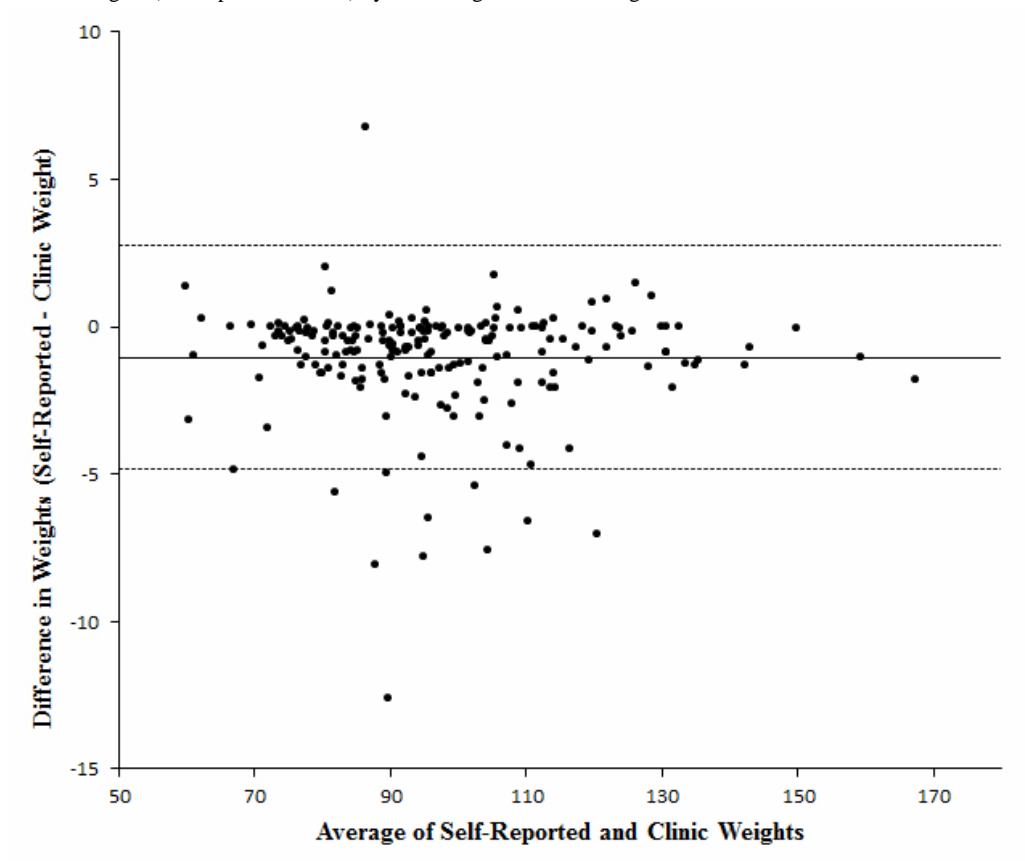
Figure 3. Accuracy of self-reported weight by percent weight change at 24 months.
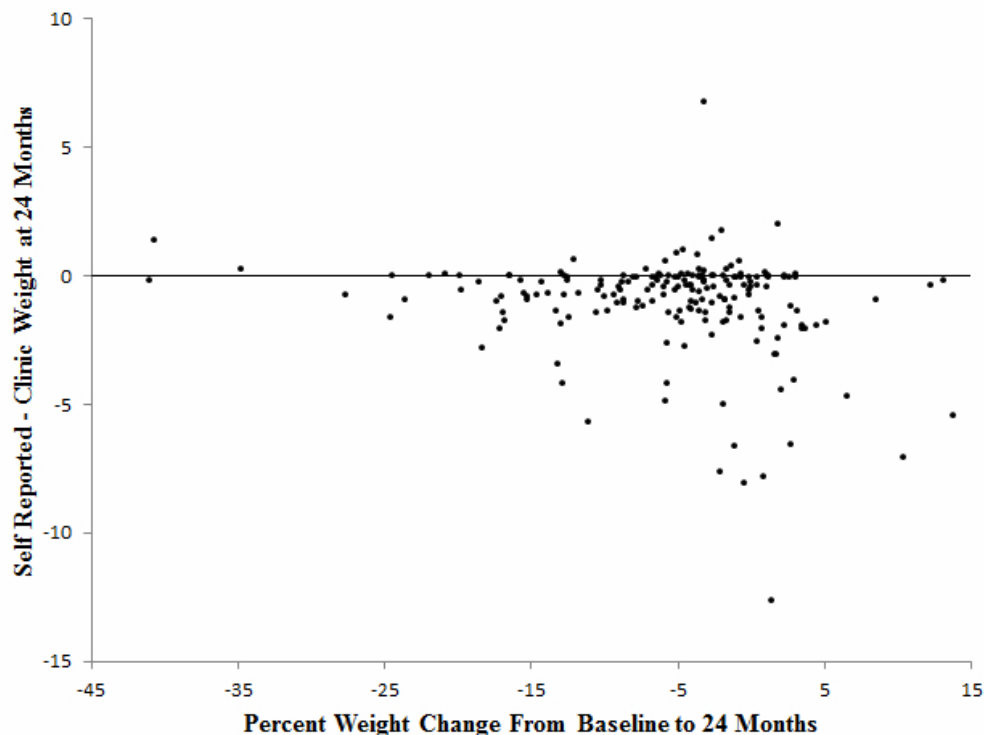
Discussion

Principal Findings

To our knowledge, this is the first study to document decreases in accuracy of self-reported weights over long-term follow-up. Specifically, participants' self-report resulted in a modest, yet significant underestimation of weight consistent with previous findings from cross-sectional studies and weight-loss trials [4,5,9,10]. The magnitude of underestimation doubled between Months 6 and 24. Moreover, weight loss was positively associated with accuracy; those who achieved the 5% weight-loss goal had more accurate self-reports. These results are congruent with findings from a study with 6 months of follow-up [10]. Strengths of the current analyses include a relatively large sample of obese adults with gender and ethnic diversity and a long duration of follow-up (ie, 24 months).

There are a number of possible sources of variation between self-reported weight and clinic weights. There was likely a difference in the accuracy of the scales provided to participants and the clinic scale that had annual accuracy verification. Another possible source of variation was diurnal weight fluctuations and other factors associated with the time lag between weights (eg, actual weight change, menstrual cycle). However, the regression analyses found a significant association among absolute difference and weight change even after controlling for length of time between measures. Although the participants were encouraged to weigh themselves under the same conditions experienced in the clinic, there may have been differences in procedures (eg, amount of clothing, placing scale on hard surface) associated with differences.

One alternative to manually entering self-reported weights on a website is the use of scales with wireless connections to a computer that can automatically upload data. However, little is known about the accuracy of this method and sources of variation may be similar to those outlined above. Pronk et al demonstrated that participants can improve the accuracy of self-reported weights with regular feedback [13]. Improvements in the accuracy of self-reporting may help individuals with long-term adherence to lifestyle-based programs.

Although short-term fluctuation in weight is expected, the extent to which small changes and the tracking of subtle trends impacts weight loss is unclear. During weight-loss maintenance, it can be difficult to reverse the trajectory of even minor weight gains [14]. Less accurate self-assessment could create challenges in identifying small changes. It is possible that the tendency to underestimate weight and increased error over time create additional barriers to long-term weight-loss efforts.

The underestimation of clinic weights in the current study was modest, and the sensitivity and specificity appears acceptable. Other studies have reported strong correlations between self-reported and clinic weights and describe the differences between self-reported and clinic weights as relatively small [9,13]. However, researchers should be cautious in using self-reported weight as a clinical outcome. The current results are consistent with a previous report of modest but significant differences between weight change calculated with self-reported versus clinic values [10]. Moreover, the generalizability of the findings are limited by the completer analysis. In the current study, 95% of the sample had 24-month clinic weights, yet only 69% had both self-reported and clinic weights. Those with self-reported weight had three times the weight loss compared to those without self-reported weights at 24 months, and this is concordant with previous reports [10].

Conclusions

Although it is not clear if accurate self-weighing facilitated weight loss or if weight loss encouraged more exacting self-assessment, those in weight-loss programs should be aware of the tendency for decreased accuracy of self-reported weight over time and the association of accurate self-report with achievement of weight-loss goals. Increased emphasis on accurate self-weighing may be an important addition to lifestyle-based weight-loss programs.

Conflicts of Interest

Healthways, Inc developed the website for both interventions used in the POWER trial in collaboration with Johns Hopkins investigators and provided coaching effort for the remotely delivered intervention. Healthways also provided some research funding to supplement National Institutes of Health support. Under an institutional consulting agreement with Healthways, the Johns Hopkins University received fees for advisory services to Healthways during the POWER trial. Faculty members who participated in the consulting services received a portion of the university fees.

On the basis of POWER trial results, Healthways developed and is commercializing a weight-loss intervention program called *Innergy*. Under an agreement with Healthways, Johns Hopkins faculty monitor the Innergy program's content and process (staffing, training, and counseling) and outcomes (engagement and weight loss) to ensure consistency with the corresponding arm of the POWER Trial. Johns Hopkins receives fees for these services and faculty members who participate in the consulting services receive a portion of these fees. Johns Hopkins receives a royalty on sales of the Innergy program.

References

1. Tufano JT, Karras BT. Mobile eHealth interventions for obesity: a timely opportunity to leverage convergence trends. J Med Internet Res 2005;7(5):e58 [FREE Full text] [doi: 10.2196/jmir.7.5.e58] [Medline: 16403722]
2. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
3. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
4. Bonn SE, Surkan PJ, Lagerros YT, Bälter K. Feasibility of a novel web-based physical activity questionnaire for young children. Pediatr Rep 2012 Dec 6;4(4):e37 [FREE Full text] [doi: 10.4081/pr.2012.e37] [Medline: 23355937]
5. Pursey K, Burrows TL, Stanwell P, Collins CE. How accurate is web-based self-reported height, weight, and body mass index in young adults? J Med Internet Res 2014;16(1):e4 [FREE Full text] [doi: 10.2196/jmir.2909] [Medline: 24398335]
6. Merrill RM, Richardson JS. Validity of self-reported height, weight, and body mass index: findings from the National Health and Nutrition Examination Survey, 2001-2006. Prev Chronic Dis 2009 Oct;6(4):A121 [FREE Full text] [Medline: 19754997]
7. Gillum RF, Sempos CT. Ethnic variation in validity of classification of overweight and obesity using self-reported weight and height in American women and men: the third National Health and Nutrition Examination Survey. Nutr J 2005;4:27 [FREE Full text] [doi: 10.1186/1475-2891-4-27] [Medline: 16209706]
8. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. Obes Rev 2007 Jul;8(4):307-326. [doi: 10.1111/j.1467-789X.2007.00347.x] [Medline: 17578381]
9. Mattila E, Lappalainen R, Pärkkä J, Salminen J, Korhonen I. Use of a mobile phone diary for observing weight management and related behaviours. J Telemed Telecare 2010;16(5):260-264. [doi: 10.1258/jtt.2009.091103] [Medline: 20483880]
10. Harvey-Berino J, Krukowski RA, Buzzell P, Ogden D, Skelly J, West DS. The accuracy of weight reported in a web-based obesity treatment program. Telemed J E Health 2011 Nov;17(9):696-699 [FREE Full text] [doi: 10.1089/tmj.2011.0032] [Medline: 21882997]
11. Appel LJ, Clark JM, Yeh HC, Wang NY, Coughlin JW, Daumit G, et al. Comparative effectiveness of weight-loss interventions in clinical practice. N Engl J Med 2011 Nov 24;365(21):1959-1968 [FREE Full text] [doi: 10.1056/NEJMoa1108660] [Medline: 22085317]
12. Jerome GJ, Yeh H, Dalcin A, Reynolds J, Gauvey-Kern ME, Charleston J, et al. Treatment of obesity in primary care practice: The Practice Based Opportunities for Weight Reduction (POWER) Trial at Johns Hopkins. Obesity and Weight Management 2009 Oct;5(5):216-221. [doi: 10.1089/obe.2009.0506]
13. Pronk NP, Crain AL, Vanwormer JJ, Martinson BC, Boucher JL, Cosentino DL. The use of telehealth technology in assessing the accuracy of self-reported weight and the impact of a daily immediate-feedback intervention among obese employees. Int J Telemed Appl 2011;2011:909248 [FREE Full text] [doi: 10.1155/2011/909248] [Medline: 21760782]

14.   Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. N Engl J
      Med 2006 Oct 12;355(15):1563-1571. [doi: 10.1056/NEJMoa061883] [Medline: 17035649]

Abbreviations

BMI: body mass index
POWER: Practice-Based Opportunities for Weight Reduction trial

XSL•FO
RenderX

Original Paper

Predictors of eHealth Usage: Insights on The Digital Divide From the Health Information National Trends Survey 2012

Emily Kontos[1], ScM, ScD; Kelly D Blake[2], ScD; Wen-Ying Sylvia Chou[2], PhD; Abby Prestin[3], PhD

[1]Harvard School of Public Health, Department of Social and Behavioral Sciences, Boston, MA, United States

[2]National Cancer Institute, Division of Cancer Control and Population Sciences, Bethesda, MD, United States

[3]United States Food and Drug Administration, Center for Tobacco Products, Office of Science, Silver Spring, MD, United States

Corresponding Author:
Emily Kontos, ScM, ScD
Harvard School of Public Health
Department of Social and Behavioral Sciences
401 Park Drive
Boston, MA, 02115
United States
Phone: 1 617 384 8724
Fax: 1 617 384 8724
Email: emily_kontos@dfci.harvard.edu

Abstract

Background: Recent eHealth developments have elevated the importance of assessing the extent to which technology has empowered patients and improved health, particularly among the most vulnerable populations. With noted disparities across racial and social groups in chronic health outcomes, such as cancer, obesity, and diabetes, it is essential that researchers examine any differences in the implementation, uptake, and impact of eHealth strategies across groups that bear a disproportionate burden of disease.

Objective: The goal was to examine eHealth use by sociodemographic factors, such as race/ethnicity, socioeconomic status (SES), age, and sex.

Methods: We drew data from National Cancer Institute's 2012 Health Information National Trends Survey (HINTS) (N=3959) which is publicly available online. We estimated multivariable logistic regression models to assess sociodemographic predictors of eHealth use among adult Internet users (N=2358) across 3 health communication domains (health care, health information–seeking, and user-generated content/sharing).

Results: Among online adults, we saw no evidence of a digital use divide by race/ethnicity. However, there were significant differences in use by SES, particularly for health care and health information–seeking items. Patients with lower levels of education had significantly lower odds of going online to look for a health care provider (high school or less: OR 0.50, 95% CI 0.33-0.76) using email or the Internet to communicate with a doctor (high school or less: OR 0.46, 95% CI 0.29-0.72), tracking their personal health information online (high school or less: OR 0.53, 95% CI 0.32-0.84), using a website to help track diet, weight, and physical activity (high school or less: OR 0.64, 95% CI 0.42-0.98; some college: OR 0.67, 95% CI 0.49-0.93), or downloading health information to a mobile device (some college: OR 0.54, 95% CI 0.33-0.89). Being female was a consistent predictor of eHealth use across health care and user-generated content/sharing domains, whereas age was primarily influential for health information–seeking.

Conclusions: This study illustrates that lower SES, older, and male online US adults were less likely to engage in a number of eHealth activities compared to their counterparts. Future studies should assess issues of health literacy and eHealth literacy and their influence on eHealth engagement across social groups. Clinical care and public health communication efforts attempting to leverage Web 2.0 and 3.0 platforms should acknowledge differential eHealth usage to better address communication inequalities and persistent disparities in health.

KEYWORDS

health communication; communication barriers; Internet; consumer health information

XSL•FO
RenderX

Introduction

The movement in the past decade toward patient-centered care has increasingly emphasized patient empowerment in health care. In particular, the Chronic Care Model, characterized by the interaction between an "informed and activated patient" and a "prepared and proactive practice team" has been highlighted as a fundamental model for optimum care [1,2]. Alongside the changing tide in health care delivery, there has been a revolution in information technology. With the development of new technology and Web 2.0 and 3.0 communication media, the field of eHealth has emerged and with it a plethora of new opportunities for individuals to access and exchange health information, manage their health through electronic platforms, and participate in "peer-to-peer health care" [3-5]. These online opportunities have been identified as a means to better enable patient empowerment and self-management of care [6].

The field of eHealth has enabled public health and medical practitioners to communicate with patients in both traditional and novel ways to address health concerns such as diabetes management [7], heart health [8], cancer prevention and health promotion activities [9-11], and smoking cessation [12,13]. Online strategies range from adaptations of more traditional communication methods, such as the delivery of tailored information and the creation of support networks [14], to more innovative developments, such as the implementation of smartphone applications for disease prevention and management [15-17].

These eHealth developments have elevated the importance of assessing the extent to which technology has empowered patients and improved health in general and among the most vulnerable populations in particular [18-21]. With noted disparities across racial and social groups in chronic health outcomes, such as cancer, obesity, and diabetes, it is essential that researchers thoughtfully examine any differences in the implementation, uptake, and impact of eHealth strategies across groups that bear a disproportionate burden of disease [20,22,23]. Current eHealth studies are limited in that many, such as those published by the Pew Internet & American Life Project, report national percentages without rigorous statistical control to determine what factors may be true drivers of any eHealth disparities. Other research has focused on issues of access to technology based on the original concept of the digital divide, which formulated 2 groups—those with access to the Internet and those without [5,24]. Gaps in access to the Internet have been persistent in that lower socioeconomic status, minority racial/ethnic groups, older age, and poorer health, among others, are associated with decreased access to the Internet [19,21,25-29]. These access patterns can lead to differential access to health information that might intensify health disparities [30]. Yet, this dichotomous oversimplification wrongly suggests that all "haves" use the Internet in a similar manner.

However, the limited number of published studies focused on use of the Internet for health point to noted communication inequalities or differences in use and engagement across important racial and social groups [18,22,31]. For example, those in the lowest income and education brackets are shown to be considerably less likely to seek out health information online compared to those in higher income and education brackets. Similarly, non-Hispanic blacks and Hispanics are significantly less likely to seek out health information online compared to their non-Hispanic white counterparts, yet these differences by race/ethnicity have begun to narrow overall. Recent unadjusted Pew Internet & American Life Project data highlight differences in topic-specific seeking behavior in that more non-Hispanic blacks and Hispanics report using the Internet to find information on how to lose weight and pregnancy compared to non-Hispanic whites, whereas a larger percentage of non-Hispanic whites report using the Internet to find information on a specific disease or problem [5,20,32,33]. These differences in eHealth use are of importance to public health practitioners and health care providers in that these communication behaviors could lead to significant health-related disparities [22].

Evidence in support of communication inequalities in health-related Internet use is building [20,34,35]; however, we lack a clear understanding of comprehensive differences by sociodemographic factors, such as race/ethnicity, socioeconomic status (SES), and sex in the utility of the Internet for eHealth tasks. Past studies have primarily examined online disparities in isolation, have not adequately adjusted for confounding factors that could drive use, or have used older datasets that do not reflect the ever-changing digital landscape [5,34,36-38].

This study aims to employ an up-to-date, comprehensive examination of eHealth use by sociodemographic factors to illustrate potential profiles of disparities across a number of communication domains. We hope this work will assist future health communication interventions and efforts that seek to use the Internet, email, and social media to reach and engage underserved populations.

Methods

Data for this study were drawn from the National Cancer Institute's 2012 Health Information National Trends Survey (HINTS). HINTS is a nationally representative survey of the US noninstitutionalized adult population that collects data on the American public's need for, access to, and use of health-related information [39]. HINTS is publicly available online [40]. Data used in this study are from HINTS 4 Cycle 1, collected from October 2011 to February 2012 (N=3959) through mailed questionnaire. The sample design was a 2-stage stratified sample with addresses selected from a comprehensive United States Postal Service national residential file, and individual respondents were selected per each household in the sample. The final response rate for HINTS 2012 was 36.7%. Further details on survey design and sampling strategies are published elsewhere [41].

To assess hypothesized differences in eHealth usage and engagement, we used 11 HINTS variables that were asked of those respondents who reported yes to ever going online to access the Internet or World Wide Web or to send and receive email (N=2358). The 11 eHealth tasks are presented in 3 domains relevant to health communication (health care, health

information-seeking, and user-generated content/sharing). Items were grouped into domains in effort to illustrate trends across eHealth tasks and for purposes of informing future health communication-related interventions [42]. The categorization of items into domains was informed by both mass communication theory, such as uses and gratifications theory, as well as recent health care policies, specifically the Affordable Care Act and Healthy People 2020, in which there is interest to track progress in goal achievement [43-45]. For example, one of the goals outlined in Healthy People 2020 is aimed at improving access to comprehensive, quality health care services, research is emerging that correlates increased engagement with the Internet and access to health care services [46]. The eHealth items assessed in this study were "In the past 12 months, have you used the Internet to look for health or medical information for yourself?" (yes/no) and "In the past 12 months, have you used the Internet for any of the following reasons?" (yes/no) as listed in Figure 1.

For the purpose of this analysis, primary predictor variables included in each model represent sociodemographic characteristics: place of birth, race/ethnicity, home ownership,

education, income, age, and sex. Hot-deck imputation was used to replace missing responses with imputed data for the race and ethnicity variables in HINTS 2012. With this approach, the resulting distribution preserved the distribution of values observed for respondents.

All models adjusted for occupational status, marital status, children, health information–seeking (ever sought health information from any source), regular access to a health care provider, insurance status, health status, personal cancer history, and family cancer history.

We used multivariable logistic regression to model the fitted odds that SES (education and income), race/ethnicity, age, and sex independently and differentially predicted eHealth usage in the population of online US adults. We used SAS-callable SUDAAN 10.0.1 to account for the complex sampling design used in HINTS and to incorporate jackknife replicate weights needed to compute accurate standard errors. All analyses were weighted to provide nationally representative estimates. We calculated weighted percentages, odds ratios (OR), and 95% confidence intervals (CI) utilizing complete case analyses with listwise deletion for each model (N=2358).

Figure 1. Health communication domains.
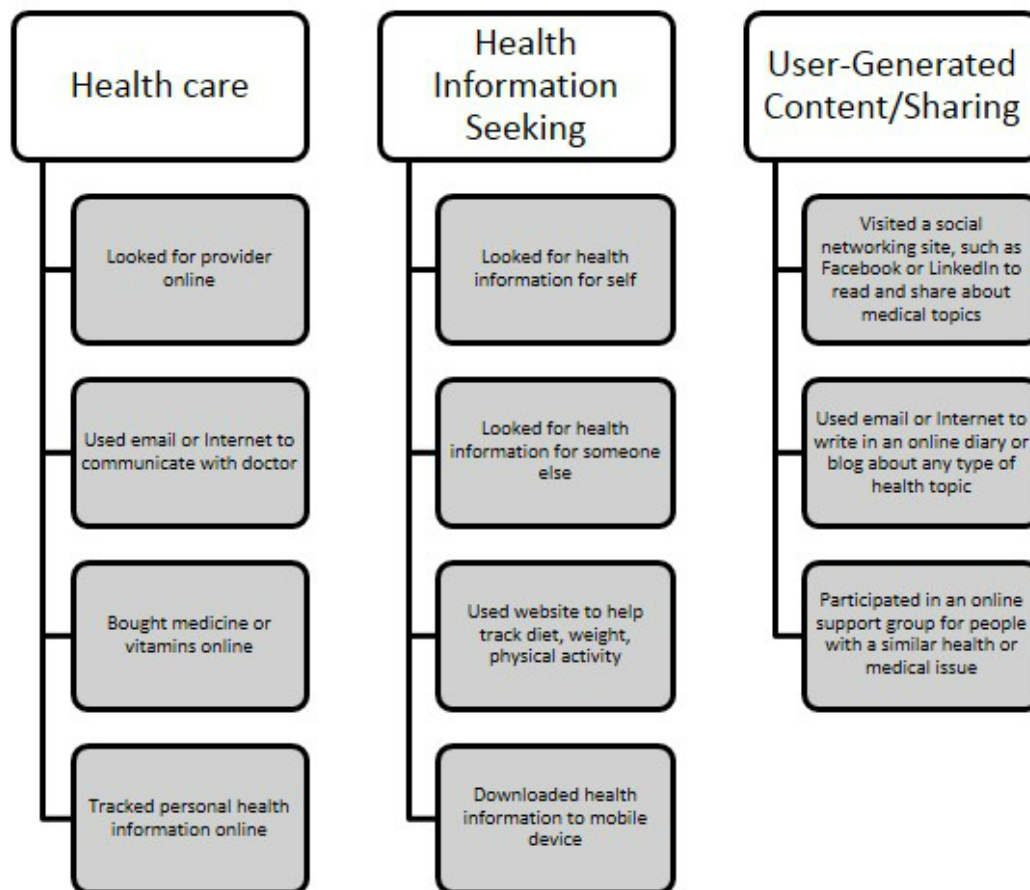
Results

Summary

Weighted and unweighted unadjusted prevalence estimates for each dependent variable are presented in Table 1 (demographics) and Table 2 (eHealth tasks). For health care–related tasks, the prevalence of eHealth usage is generally low, with approximately 18.95% (509/2358) of online US adults reporting ever having engaged in activities such as emailing providers, 19.29% (501/2358) tracking health information online, and 17.67% (459/2358) buying medicine online. Slightly more people have used the Internet to search for a health care provider (38.42%, 861/2358). For health information–seeking tasks, eHealth usage is notably more prevalent. Nearly 80% (79.04%, 1833/2358) of online American adults have used the Internet to look for health information for themselves and 57.04% (1342/2358) have used the Internet to look for health information for someone else. Approximately 42.98% (925/2358) have used the Internet in the past year to help with diet, weight, or physical activity, but far fewer have used it to download health information to a mobile device (11.70%, 261/2358). In terms of engagement with user-generated content, only a small proportion of the population (3.26%-4.63%, 76-110/2358 of online US adults) took advantage of interactive Web features, such as participating in an online support group or health-related blog. Use of social networking sites (SNS) for health is a bit higher, with 16.80% (345/2358) of online Americans saying they have visited sites such as Facebook or LinkedIn to read or share about medical topics.

Pursuant to our central research question on potential communication inequalities in eHealth usage, results for our multivariable logistic regression analyses are presented by the domains of health care, health information–seeking, and user-generated content/sharing. Among online adults, there was little evidence of a digital use divide by race/ethnicity. Generally, non-Hispanic blacks, Hispanics, and people of other races were no more or less likely than non-Hispanic whites to engage in eHealth activities. Being of other race was predictive of almost a 3-fold increase in the odds of downloading health information to a mobile device and twice the odds of using the Internet to look for a health care provider. However, the most significant differences in eHealth use were across SES (either by education, income, or both) and by age and sex. Findings are summarized subsequently.

XSL·FO
RenderX

Table 1. Weighted and unweighted unadjusted prevalence estimates for sample demographics, Health Information National Trends Survey (HINTS) 4 Cycle 1, October 2011 to February 2012 (N=2358).

Sociodemographics	n	Unweighted %	Weighted %
Age group			
18-34	440	18.66	35.21
35-49	680	28.84	30.07
50-64	824	34.94	24.63
65-74	292	12.38	7.03
>75	122	5.17	3.05
Highest level of school completed			
≤High school	401	17.01	23.58
Some college	761	32.27	35.25
College graduate or more	1196	50.72	41.17
Race/ethnicity (imputed)			
Non-Hispanic white	1657	70.27	70.43
Hispanic	215	9.12	11.50
Non-Hispanic black	322	13.66	10.39
Non-Hispanic other	164	6.96	7.69
Annual household income			
Less than $20,000	329	13.95	16.79
$20,000 to <$35,000	313	13.27	14.33
$35,000 to <$50,000	313	13.27	11.27
$50,000 to <$75,000	437	18.53	18.69
>$75,000	821	34.82	33.04
Missing	145	6.15	5.88
Sex			
Male	931	39.48	48.36
Female	1427	60.52	51.64
Ever diagnosed as having cancer?			
Yes	311	13.19	7.34
No	2047	86.81	92.66
Any family members ever had cancer?			
Yes	1506	63.87	62.92
No/not sure/missing	852	36.13	37.08

Table 2. Weighted and unweighted unadjusted prevalence estimates for eHealth tasks, Health Information National Trends Survey (HINTS) 4 Cycle 1, October 2011 to February 2012 (2012) (N=2358).

eHealth Task	n	Unweighted %	Weighted %
Ever looked for information about health or medical topics from any source?			
Yes	2132	90.42	88.79
No	226	9.58	11.21
In the past 12 months, used email or Internet to communicate with a doctor or doctor's office?			
No	1849	78.41	81.05
Yes	509	21.59	18.95
In the past 12 months, bought medicine or vitamins online?			
No	1899	80.53	82.33
Yes	459	19.47	17.67
In the past 12 months, used the Internet to look for a health care provider?			
No	1497	63.49	61.58
Yes	861	36.51	38.42
In the last 12 months, used the Internet to keep track of personal health information			
No	1857	78.75	80.71
Yes	501	21.25	19.29
In the past 12 months, used a website to help with diet, weight, or physical activity?			
No	1433	60.77	57.02
Yes	925	39.23	42.98
In the last 12 months, used the Internet to download health-related info to a mobile device			
No	2097	88.93	88.30
Yes	261	11.07	11.70
In the past 12 months, used the Internet to look for health or medical information for self?			
No	525	22.3	21
Yes	1833	77.7	79
In the past 12 months, used the Internet to look for health or medical information for someone else?			
No	1016	43.09	42.96
Yes	1342	56.91	57.04
In the past 12 months, participated in an online support group for people with a similar health or medical issue?			
No	2248	95.34	95.37
Yes	110	4.66	4.63
In the past 12 months, visited a social networking site to read and share about medical topics?			
No	2013	85.37	83.20
Yes	345	14.63	16.80
In the past 12 months, wrote in an online diary or blog about any type of health topic?			
No	2282	96.78	96.74
Yes	76	3.22	3.26

XSL•FO
RenderX

Health Care

Online adults with the lowest levels of education were significantly less likely to use the Internet to look for a health care provider compared to those with a college degree or more (OR 0.50, 95% CI 0.33-0.76) (Table 3). The youngest adults surveyed (18-34 years) had more than twice the odds of engaging in online provider searches compared to the oldest group aged 65 years and older (OR 2.24, 95% CI 1.20-4.16). Whereas, women were more likely than men to search for a health care provider online (OR 1.53, 95% CI 1.14-2.04).

Among online adults, we again saw evidence of a usage gap by education and sex with regard to using email or the Internet to communicate with a doctor or doctor's office. Those with a high school degree or less were less likely to engage in this activity compared to those with a college degree or more (OR 0.46, 95% CI 0.29-0.72). Additionally, women were more likely than men to have communicated with a provider by email or Internet (OR 1.52, 95% CI 1.06-2.19).

Education and sex were also significant predictors of tracking personal health information online. High school graduates and those with lower levels of education were less likely than college graduates to have tracked this information online (OR 0.53, 95% CI 0.32-0.84), whereas women were 1.5 times as likely to have done so than men were (OR 1.52, 95% CI 1.06-2.19).

We saw different patterns in use when examining who purchased medicine or vitamins online. Income level and place of birth were significant predictors; those earning <$20,000 and between $20,000 and <$35,000 annually were significantly less likely than those in the highest income category to have made medicine or vitamin purchases online (OR 0.34, 95% CI 0.12-0.95 and OR 0.38, 95% CI 0.16-0.90, respectively), whereas those not born in the United States had more than 2.5 times the odds of having done so (OR 2.64, 95% CI 1.37-5.09).

Table 3. Multivariable logistic regression models for odds[a] of reporting yes to eHealth usage by socioeconomic status and race/ethnicity, Health Information National Trends Survey (HINTS), 2012.

Sociodemographics	Looked for health care provider		Used email or Internet to communicate with doctor		Bought medicine or vitamins online		Tracked personal health information online	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Education								
≤High school degree	0.50	0.33, 0.76	0.46	0.29, 0.72	0.67	0.34, 1.32	0.53	0.32, 0.84
Some college	0.73	0.49, 1.09	0.71	0.50, 1.02	0.99	0.67, 1.46	0.75	0.50, 1.13
>College degree (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Born in United States								
No	1.29	0.73, 2.26	1.03	0.49, 2.14	2.64	1.37, 5.09	1.19	0.58, 2.46
Yes (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Race/ethnicity								
Hispanic	0.73	0.44, 1.16	1.21	0.69, 2.13	1.01	0.49, 2.10	1.04	0.60, 1.81
Non-Hispanic black	1.52	0.72, 3.18	0.94	0.49, 1.79	1.39	0.50, 3.91	1.23	0.73, 2.06
Other race	2.02	1.01, 4.04	1.68	0.91, 3.11	0.79	0.43, 1.46	1.82	0.81, 4.07
Non-Hispanic white (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Home ownership								
Rent or occupy without rent	1.06	0.67, 1.68	1.04	0.72, 1.51	0.81	0.51, 1.28	1.19	0.86, 1.64
Own (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Annual household income								
<$20,000	0.86	0.38, 1.97	0.37	0.13, 1.02	0.34	0.12, 0.95	0.78	0.31, 1.94
$20,000 to <$35,000	0.70	0.35, 1.38	0.49	0.23, 1.06	0.38	0.16, 0.90	0.81	0.45, 1.46
$35,000 to <$50,000	1.51	0.95, 2.40	0.67	0.40, 1.13	0.70	0.32, 1.54	0.64	0.36, 1.15
$50,000 to <$75,000	1.00	0.64, 1.58	0.71	0.47, 1.06	0.91	0.60, 1.37	1.31	0.80, 2.14
>$75,000 (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Age								
18-34	2.24	1.20, 4.16	0.91	0.48, 1.70	0.85	0.46, 1.58	1.24	0.69, 2.22
35-49	1.55	0.91, 2.63	0.77	0.47, 1.24	0.68	0.40, 1.14	0.90	0.58, 1.39
50-64	1.50	0.90, 2.51	0.84	0.51, 1.37	0.73	0.46, 1.16	0.91	0.65, 1.28
>65 (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Sex								
Female	1.53	1.14, 2.04	1.47	1.03, 2.09	0.98	0.65, 1.47	1.52	1.06, 2.19
Male (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00

[a]All estimates are weighted. All models control for occupational status, marital status, children, health information–seeking (ever sought health information from any source), regular access to a health care provider, insurance status, health status, personal cancer history, and family history of cancer.

Health Information–Seeking

Age was the sole predictor of whether online adults used the Internet in the past 12 months to search for health information for themselves (Table 4). Adults aged 18-34 years were 3.5 times as likely and adults aged 35-49 years were nearly 2.5 times as likely as those 65 years and older to use the Internet to search for health information (OR 3.51, 95% CI 1.66-7.44 and OR 2.35, 95% CI 1.17-4.72, respectively).

Age and sex were differentially predictive of using the Internet to search for health or medical information for someone else.

Those aged 35-49 years were 1.5 times as likely to have used the Internet for this purpose as those 65 years and older (OR 1.52, 95% CI 1.00-2.31). Women were approximately 1.5 times as likely as men to have done so (OR 1.46, 95% CI 1.06-2.01).

Similarly, we identified gaps by age, education, and income in use of websites to help with diet, weight, or physical activity. Overwhelmingly, those in younger age categories were significantly more likely than those aged 65 and older to have used a website for this purpose: age 18-34 (OR 3.37, 95% CI 2.00-5.69), age 35-49 (OR 2.57, 95% CI 1.66, 3.99), and age 50-64 (OR 2.22, 95% 1.43-3.43). Those with a high school

XSL·FO
RenderX

degree or less and those with some college were approximately 35% less likely than college graduates to have done so (OR 0.64, 95% CI 0.42-0.98 and OR 0.67, 95% CI 0.49-0.93, respectively). Additionally, those making less than $20,000 per year were nearly 50% less likely than those in the highest income category to have used the Web for this purpose (OR 0.46, 95% CI 0.24-0.86).

In terms of downloading health-related information to a mobile device, such as an MP3 player, cell phone, tablet computer, or electronic book device, we observed a few usage gaps by education level and race/ethnicity. Those with some college were 46% less likely than those with a college degree to have gone online for this purpose (OR 0.54, 95% CI 0.33-0.89) and those in the lowest education bracket were 58% less likely to engage in this eHealth task, but this did not meet statistical significance (OR 0.42, 95% CI 0.17-1.02). Those of other race were nearly 3 times as likely to download health information to a mobile device compared to their non-Hispanic white counterparts (OR 2.78, 95% CI 1.33-5.86).

Table 4. Multivariable logistic regression models for odds[a] of reporting yes to eHealth usage, by socioeconomic status and race/ethnicity, Health Information National Trends Survey (HINTS), 2012.

Sociodemographics	Looked for health information for self		Looked for health information for someone else		Used website to help track diet, weight, physical activity		Downloaded health information to mobile device	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Education								
≤High school degree	0.64	0.41, 1.01	0.72	0.46, 1.10	0.64	0.42, 0.98	0.42	0.17, 1.02
Some college	0.67	0.44, 1.02	0.70	0.46, 1.05	0.67	0.49, 0.93	0.54	0.33, 0.89
>College degree (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Born in United States								
No	1.43	0.68, 2.99	1.29	0.67, 2.48	1.39	0.77, 2.49	1.36	0.65, 2.88
Yes (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Race/ethnicity								
Hispanic	0.62	0.27, 1.45	0.89	0.51, 1.57	1.58	0.80, 3.11	1.47	0.57, 3.82
Non-Hispanic black	1.21	0.66, 2.23	1.32	0.65, 2.68	1.00	0.55, 1.82	1.73	0.88, 3.42
Other race	1.13	0.40, 3.20	1.78	0.94, 3.34	1.40	0.75, 2.60	2.78	1.33, 5.86
Non-Hispanic white (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Home ownership								
Rent or occupy without rent	0.89	0.59, 1.33	1.13	0.72, 1.79	1.06	0.72, 1.57	1.18	0.62, 2.24
Own (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Annual household income								
<$20,000	0.84	0.39, 1.80	1.06	0.52, 2.15	0.46	0.24, 0.86	1.28	0.59, 2.80
$20,000 to <$35,000	0.83	0.40, 1.74	1.45	0.72, 2.90	0.72	0.39, 1.32	0.88	0.37, 2.13
$35,000 to <$50,000	0.94	0.50, 1.75	1.10	0.60, 2.00	0.59	0.34, 1.05	0.81	0.38, 1.75
$50,000 to <$75,000	0.87	0.47, 1.59	1.02	0.68, 1.52	1.12	0.74, 1.72	1.09	0.61, 1.96
>$75,000 (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Age								
18-34	3.51	1.66, 7.44	1.31	0.74, 2.31	3.37	2.00, 5.69	1.78	0.77, 4.09
35-49	2.35	1.17, 4.72	1.52	1.00, 2.31	2.57	1.66, 3.99	0.89	0.43, 1.84
50-64	1.62	0.87, 3.02	1.44	0.95, 2.19	2.22	1.43, 3.43	1.32	0.66, 2.62
>65 (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Gender								
Female	1.43	0.96, 2.12	1.46	1.06, 2.01	1.25	0.93, 1.70	1.21	0.75, 1.93
Male (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00

[a]All estimates are weighted. All models control for occupational status, marital status, children, health information–seeking (ever sought health information from any source), regular access to a health care provider, insurance status, health status, personal cancer history, and family history of cancer.

XSL•FO
RenderX

Engagement in User-Generated Content and Social Media

We saw differences in health-related social media use by SES, sex, and age among online adults (Table 5). Both lower education and lower income were predictive of using SNS, such as Facebook, to read or share about medical topics. Those with some college were more than 1.5 times as likely as those with a college degree to engage in this eHealth activity (OR 1.59, 95% CI 1.06-2.39). Similarly, those with household incomes less than $20,000 were more than twice as likely to use SNS to read or share about health (OR 2.12, 95% CI 1.04-4.29) and those earning $20,000-$35,000 were also nearly twice as likely to engage in this activity compared to those with a household income of $75,000 or more, although this did not meet statistical significance (OR 1.82, 95% CI 0.99-3.32). Moreover, we saw a fine gradation of effect from the youngest to oldest age categories for SNS health use. Compared to those aged 65 and older, those aged 18-34 were nearly 3 times more likely to have gone online for this purpose and those aged 35-49 were more than twice as likely to have done so (OR 2.81, 95% CI 1.13-7.00 and OR 2.27, 95% CI 1.00-5.17, respectively).

Participation in an online support group for people with a similar medical issue was more prominent among women; women were nearly 3 times as likely to participate in an online support group compared to men (OR 2.79, 95% CI 1.20-6.51). Those with a household income of $50,000-$75,000 were more than twice as likely compared to the highest income bracket to do so (OR 2.22, 95% CI 1.04-4.73). In terms of writing an online blog about a health topic, women were more than 4 times as likely as men to have done so (OR 4.31, 95% CI 1.78-10.42).

Having a connection to cancer was also predictive of engaging in user-generated content for health. Respondents with a family history of cancer were nearly 3 times as likely to have participated in an online support group than those without a family cancer experience (OR 2.96, 95% CI 1.00-3.83). Cancer survivors were nearly 3 times as likely to have written in a blog about a health topic compared to those with no cancer history (OR 2.93, 95% CI 1.00-8.63).

Table 5. Multivariable logistic regression models for odds[a] of reporting yes to eHealth usage, by SES and race/ethnicity, Health Information National Trends Survey (HINTS), 2012.

Sociodemographics	Visited a social networking site to read and share about medical topics		Used email or Internet to write in an online diary or blog about any type of health topic		Participated in an online support group for people with a similar health or medical issue[b]	
	OR	95% CI	OR	95% CI	OR	95% CI
Education						
≤High school	1.11	0.64, 1.92	1.05	0.38, 2.90	0.90	0.27, 3.00
Some college	1.59	1.06, 2.39	1.13	0.45, 2.85	1.44	0.79, 2.64
College degree or more (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Born in United States						
No	0.77	0.42, 1.39	1.43	0.31, 6.68	1.96	0.36, 8.86
Yes (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Race/ethnicity						
Hispanic	0.90	0.39, 2.09	0.62	0.12, 3.20	1.91	0.61, 5.95
Non-Hispanic black	1.24	0.57, 2.67	1.32	0.39, 4.53	0.57	0.17, 1.88
Other race	1.08	0.48, 2.43	1.33	0.26, 6.83	1.80	0.50, 6.45
Non-Hispanic white (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Home ownership						
Rent or occupy without rent	1.23	0.72, 2.10	1.18	0.37, 3.71	0.83	0.34, 2.01
Own (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Annual household income						
<$20,000	2.12	1.04, 4.29	0.87	0.24, 3.15	0.87	0.13, 5.83
$20,000 to <$35,000	1.82	0.99, 3.32	2.41	0.73, 7.95	1.31	0.33, 5.26
$35,000 to <$50,000	1.56	0.82, 2.94	1.18	0.40, 3.55	1.02	0.28, 3.79
$50,000 to <$75,000	1.12	0.59, 2.13	1.28	0.50, 3.28	2.22	1.04, 4.73
>$75,000 (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Age						
18-34	2.81	1.13, 7.00	3.91	0.59, 25.94	1.99	0.47, 8.45
35-49	2.27	1.00, 5.17	1.65	0.27, 10.23	1.38	0.34, 5.65
50-64	1.59	0.72, 3.49	1.70	0.25, 11.68	0.86	0.15, 4.97
>65 (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Sex						
Female	1.48	0.88, 2.49	4.31	1.78, 10.42	2.79	1.20, 6.51
Male (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Cancer experience						
Cancer diagnosis (self)	1.07	0.63, 1.81	2.93	1.00, 8.63	1.36	0.62, 2.99
No cancer diagnosis (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00
Family history of cancer	1.08	0.72, 1.62	1.14	0.52, 2.53	2.96	1.00, 3.83
No family history (ref)	1.00	1.00, 1.00	1.00	1.00, 1.00	1.00	1.00, 1.00

[a]All estimates are weighted. All models control for occupational status, marital status, children, health information–seeking (ever sought health information from any source), regular access to a health care provider, insurance status, health status, personal cancer history, and family history of cancer.

[b]For the online support group model, high school degree and no high school degree were collapsed into 1 predictor variable to increase cell size for analysis.

XSL•FO
RenderX

Discussion

Being younger and female has consistently been predictive of increased use of eHealth [5,21,25,47]. Younger generations who have grown up with technology have been labeled "digital natives" and they are more comfortable using technology for everyday needs, including management of their health care needs. In comparison, older generations, labeled "digital immigrants," have had to learn and acquire the necessary skills needed to navigate the Internet and are generally less comfortable using technology [48]. Females also tend to have increased eHealth utilization due in part to their higher engagement in both health care-related online activities and increased use of general social media, such as SNS [3-5,49]. This could be because of their role as the health care liaison for their family members.

Our analysis identifies specific proxies of SES that are more reliably associated with eHealth use. Education was more consistently predictive of eHealth use across the health care and information-seeking domains, whereas household income was somewhat less predictive across items and domains. Although both education and income have been used to describe SES, in considering technology use and health communication, our analysis suggests that education may be a more salient proxy as compared to income.

Furthermore, this distinction between common proxies for SES may offer insight into the fundamental drivers of online use for health and subsequently assist in the development of more effective interventions and programs. Divides among those less educated indicate that issues of health literacy and eHealth literacy may be important factors. Research is emerging that advocates matching eHealth technology to the eHealth literacy (defined as "the ability to seek, find, understand, and appraise health information from electronic sources and apply knowledge gained to addressing or solving a health problem") of the intended user [50-52]. Future analyses of online communication-based interventions should better investigate and address issues pertaining to eHealth literacy in an effort to reduce communication inequalities across low-SES groups.

It is also important to examine the direction of the relationship between education and engagement in eHealth tasks. Lower levels of education were associated with increased use of social media for health, whereas higher levels of education were associated with increased use of the Internet to engage in health care-related activities and search for information—much like the patterns we see for static Web 1.0 engagement. Similar patterns are seen in the Pew Internet & American Life Project's 2013 Health Online study in that higher percentages of non-Hispanic blacks (35%) and Hispanics (38%) reported using their mobile phones to access health information compared to non-Hispanic whites (27%) [5]. These differences in use across domains may indicate potential inequalities in not only the quality of information obtained, but also patients' engagement with the health care system. Further empirical examination is warranted to better objectively assess quality of health information shared via social media compared with more basic Web 1.0 sites as well as quality and satisfaction of care between those that engage in eHealth care and those that do not.

On the other hand, our study also indicates the potential for eHealth technologies to aid in reducing communication inequalities and disparities in health. Our data found no racial divides among the most vulnerable groups in eHealth use once access was achieved; there was only increased use of downloading information to a mobile device and looking online for a health care provider among other race individuals. This finding is consistent with several prior studies [5,25,35] and points to the opportunity to better explore the association between eHealth utilization and health outcomes among Hispanic and non-Hispanic blacks more directly.

Although our research is an important addition to the literature, we note its limitations. First, the low survey response rate may increase sampling error in our estimates; however, overall sampling coverage was enhanced through the stratified design. Also, because this was a cross-sectional survey study, it is challenging to account for unmeasured confounding variables. Future studies should examine potential factors related to those sociodemographic variables that were predictive of our outcomes, such as literacy, to determine their role in eHealth use. In addition, our study does not attempt to examine factors that may moderate or perhaps mediate the main associations presented within this analysis. For example, past research has examined the role of trust in terms of patients' trust of the Internet as well as their own physicians and/or the health care system [53,54]. Examination of these psychographic variables is an important addition in better understanding the complex predictors of eHealth use. Yet, with such persistent inequalities in health, we cannot underestimate the importance of maintaining a constant understanding of eHealth use by sociodemographic characteristics.

A more qualitative examination into the rationale offered by patients for their eHealth utilization would build upon this work and offer a more robust understanding of why certain groups do and do not use the Internet for health purposes. A recent study conducted by the Pew Internet & American Life Project reports that 36% of non-Internet users cited that they did not use the Internet because they did not think that it was relevant to them, whereas 32% reported difficulty in use [55]. Investigation as to whether or not these reasons are applicable to eHealth use is warranted.

This study illustrates that lower SES, older, and male online US adults were less likely to engage in a number of eHealth activities compared to their counterparts. In an effort to reduce existing disparities in health outcomes, clinical and public health communication strategies should be attuned to these differences in online use.

Our results have important implications for clinical care and public health communication efforts attempting to leverage Web 2.0 and 3.0 platforms. It is evident that a one size/platform would not fit all, as significant demographic and individual factors influence eHealth engagement. For example, campaigns and interventions targeting women or younger populations may see success in utilizing user-generated content and sharing sites. Whereas, in the current Chronic Care Model of medical care,

offline materials should perhaps supplement online health information if practitioners would like to ensure equal access to information across educational strata, age, and sex.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Glasgow RE, Orleans CT, Wagner EH. Does the chronic care model serve also as a template for improving prevention? Milbank Q 2001;79(4):579-612, iv [FREE Full text] [Medline: 11789118]

2. Wagner EH, Bennett SM, Austin BT, Greene SM, Schaefer JK, Vonkorff M. Finding common ground: patient-centeredness and evidence-based chronic illness care. J Altern Complement Med 2005;11 Suppl 1:S7-15. [doi: 10.1089/acm.2005.11.s-7] [Medline: 16332190]

3. Fox S, Jones S. The Social Life of Health Information. Washington, DC: Pew Internet & American Life Project; 2009 May 12. URL: http://www.pewinternet.org/Reports/2011/Social-Life-of-Health-Info/Summary-of-Findings.aspx [accessed 2013-11-20] [WebCite Cache ID 6LH0FPJlu]

4. Fox S, Purcell K. Chronic Disease and the Internet. Washington, DC: Pew Internet & American Life Project; 2010 Mar 24. URL: http://www.pewinternet.org/Reports/2010/Chronic-Disease/Summary-of-Findings.aspx [accessed 2013-11-20] [WebCite Cache ID 6LH0LG64Z]

5. Fox S, Duggan M. Health Online. Washington, DC: Pew Internet & American Life; 2013 Jan 15. URL: http://www.pewinternet.org/Reports/2013/Health-online/Summary-of-Findings.aspx [accessed 2013-11-20] [WebCite Cache ID 6LH0VgIUW]

6. 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]

7. Vuong AM, Huber JC, Bolin JN, Ory MG, Moudouni DM, Helduser J, et al. Factors affecting acceptability and usability of technological approaches to diabetes self-management: a case study. Diabetes Technol Ther 2012 Dec;14(12):1178-1182 [FREE Full text] [doi: 10.1089/dia.2012.0139] [Medline: 23013155]

8. Lindsay S, Smith S, Bellaby P, Baker R. The health impact of an online heart disease support group: a comparison of moderated versus unmoderated support. Health Educ Res 2009 Aug;24(4):646-654 [FREE Full text] [doi: 10.1093/her/cyp001] [Medline: 19251770]

9. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]

10. Smith DT, Carr LJ, Dorozynski C, Gomashe C. Internet-delivered lifestyle physical activity intervention: limited inflammation and antioxidant capacity efficacy in overweight adults. J Appl Physiol (1985) 2009 Jan;106(1):49-56 [FREE Full text] [doi: 10.1152/japplphysiol.90557.2008] [Medline: 19008491]

11. Steele RM, Mummery WK, Dwyer T. A comparison of face-to-face or internet-delivered physical activity intervention on targeted determinants. Health Educ Behav 2009 Dec;36(6):1051-1064. [doi: 10.1177/1090198109335802] [Medline: 19502534]

12. Bock BC, Graham AL, Whiteley JA, Stoddard JL. A review of web-assisted tobacco interventions (WATIs). J Med Internet Res 2008;10(5):e39 [FREE Full text] [doi: 10.2196/jmir.989] [Medline: 19000979]

13. Graham AL, Cobb NK, Papandonatos GD, Moreno JL, Kang H, Tinkelman DG, et al. A randomized trial of Internet and telephone treatment for smoking cessation. Arch Intern Med 2011 Jan 10;171(1):46-53 [FREE Full text] [doi: 10.1001/archinternmed.2010.451] [Medline: 21220660]

14. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants. 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]

15. Muessig KE, Pike EC, Fowler B, LeGrand S, Parsons JT, Bull SS, et al. Putting prevention in their pockets: developing mobile phone-based HIV interventions for black men who have sex with men. AIDS Patient Care STDS 2013 Apr;27(4):211-222 [FREE Full text] [doi: 10.1089/apc.2012.0404] [Medline: 23565925]

16. Hebden L, Balestracci K, McGeechan K, Denney-Wilson E, Harris M, Bauman A, et al. 'TXT2BFiT' a mobile phone-based healthy lifestyle program for preventing unhealthy weight gain in young adults: study protocol for a randomized controlled trial. Trials 2013;14:75 [FREE Full text] [doi: 10.1186/1745-6215-14-75] [Medline: 23506013]

17. Pellegrini CA, Duncan JM, Moller AC, Buscemi J, Sularz A, DeMott A, et al. A smartphone-supported weight loss program: design of the ENGAGED randomized controlled trial. BMC Public Health 2012;12:1041 [FREE Full text] [doi: 10.1186/1471-2458-12-1041] [Medline: 23194256]

18. Gibbons MC, Fleisher L, Slamon RE, Bass S, Kandadai V, Beck JR. Exploring the potential of Web 2.0 to address health disparities. J Health Commun 2011;16 Suppl 1:77-89. [doi: 10.1080/10810730.2011.596916] [Medline: 21843097]

19. Kontos EZ, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. J Med Internet Res 2007;9(4):e31 [FREE Full text] [doi: 10.2196/jmir.9.4.e31] [Medline: 17951215]

20. Kontos EZ, Emmons KM, Puleo E, Viswanath K. Contribution of communication inequalities to disparities in human papillomavirus vaccine awareness and knowledge. Am J Public Health 2012 Oct;102(10):1911-1920. [doi: 10.2105/AJPH.2011.300435] [Medline: 22970692]

21. Kontos EZ, Emmons KM, Puleo E, Viswanath K. Communication inequalities and public health implications of adult social networking site use in the United States. J Health Commun 2010;15 Suppl 3:216-235 [FREE Full text] [doi: 10.1080/10810730.2010.522689] [Medline: 21154095]

22. Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. Am J Prev Med 2007 May;32(5 Suppl):S131-S133 [FREE Full text] [doi: 10.1016/j.amepre.2007.02.012] [Medline: 17466818]

23. Viswanath K. Public communications and its role in reducing and eliminating health disparities. In: Thomson GE, Mitchell F, Williams MB, editors. Examining the Health Disparities Research Plan of the National Institutes of Health: Unfinished Business. Washington, DC: The National Academies Press; 2006:215-253.

24. Zillien N. Digital distinction: status-specific types of Internet usage. Social Science Quarterly 2009 Apr 13;90(2):274-291. [doi: 10.1111/j.1540-6237.2009.00617.x]

25. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]

26. Jackson LA, Zhao Y, Kolenic A, Fitzgerald HE, Harold R, Von Eye A. Race, gender, and information technology use: the new digital divide. Cyberpsychol Behav 2008 Aug;11(4):437-442. [doi: 10.1089/cpb.2007.0157] [Medline: 18721092]

27. Wang JY, Bennett K, Probst J. Subdividing the digital divide: differences in internet access and use among rural residents with medical limitations. J Med Internet Res 2011;13(1):e25 [FREE Full text] [doi: 10.2196/jmir.1534] [Medline: 21371989]

28. Jansen J. Use of Internet in higher-income households. Washington, DC: Pew Internet & American Life Project; 2010 Nov 24. URL: http://pewinternet.org/Reports/2010/Better-off-households/Overview.aspx [accessed 2013-11-20] [WebCite Cache ID 6LH1HAXJm]

29. Peterson A. Washington Post. 2013 Nov 12. Gap between those who use Internet and those who don't is widening URL: http://www.washingtonpost.com/business/technology/gap-between-those-who-use-internet-and-those-who-dont-is-widening/2013/11/12/d9d8d002-4726-11e3-a196-3544a03c2351_story.html [accessed 2013-11-20] [WebCite Cache ID 6LH1LatXt]

30. Richardson LD, Norris M. Access to health and health care: how race and ethnicity matter. Mt Sinai J Med 2010;77(2):166-177. [doi: 10.1002/msj.20174] [Medline: 20309927]

31. Zach L, Dalrymple PW, Rogers ML, Williver-Farr H. Assessing internet access and use in a medically underserved population: implications for providing enhanced health information services. Health Info Libr J 2012 Mar;29(1):61-71. [doi: 10.1111/j.1471-1842.2011.00971.x] [Medline: 22335290]

32. Kontos EZ, Viswanath K. Cancer-related direct-to-consumer advertising: a critical review. Nat Rev Cancer 2011 Feb;11(2):142-150. [doi: 10.1038/nrc2999] [Medline: 21258398]

33. Tu HT, Cohen GR. Striking jump in consumers seeking health care information. Track Rep 2008 Aug(20):1-8. [Medline: 18770913]

34. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]

35. Thackeray R, Crookston BT, West JH. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]

36. Nguyen GT, Bellamy SL. Cancer information seeking preferences and experiences: disparities between Asian Americans and Whites in the Health Information National Trends Survey (HINTS). J Health Commun 2006;11 Suppl 1:173-180. [doi: 10.1080/10810730600639620] [Medline: 16641082]

37. Cheong PH. Health communication resources for uninsured and insured Hispanics. Health Commun 2007;21(2):153-163. [doi: 10.1080/10410230701307188] [Medline: 17523861]

38. Beckjord EB, Finney Rutten LJ, Squiers L, Arora NK, Volckmann L, Moser RP, et al. Use of the internet to communicate with health care providers in the United States: estimates from the 2003 and 2005 Health Information National Trends Surveys (HINTS). J Med Internet Res 2007;9(3):e20 [FREE Full text] [doi: 10.2196/jmir.9.3.e20] [Medline: 17627929]

39. Nelson DE, Kreps GL, Hesse BW, Croyle RT, Willis G, Arora NK, et al. The Health Information National Trends Survey (HINTS): development, design, and dissemination. J Health Commun 2004;9(5):443-60; discussion 81. [doi: 10.1080/10810730490504233] [Medline: 15513791]

40.   National Cancer Insititute. 2012. Heath Information National Trends Survey URL: http://hints.cancer.gov/ [accessed 2014-06-19] [WebCite Cache ID 6QRmhWxvC]

41.   Finney Rutten LJ, Davis T, Beckjord EB, Blake K, Moser RP, Hesse BW. Picking up the pace: changes in method and frame for the health information national trends survey (2011-2014). J Health Commun 2012;17(8):979-989. [doi: 10.1080/10810730.2012.700998] [Medline: 23020763]

42.   Southwell BG, Hamilton JT, Slater JS. Why addressing the poor and underinsured is vexing. Health Commun 2011 Sep;26(6):583-585. [doi: 10.1080/10410236.2011.575453] [Medline: 21843099]

43.   Blumler JG, Katz E. The uses of mass communications: Current perspectives on gratifications research. American Journal of Sociology 1974 May;81(6):1546-1548.

44.   US Department of Health and Human Services. 2010. The Affordable Care Act URL: http://www.hhs.gov/healthcare/rights/law/index.html [accessed 2014-02-05] [WebCite Cache ID 6NABIdLaJ]

45.   HealthPeople.gov. 2013. Access to Health Services: Objectives URL: http://www.healthypeople.gov/2020/topicsobjectives2020/overview.aspx?topicid=1 [accessed 2014-02-05] [WebCite Cache ID 6NABSNDwC]

46.   Lee CJ, Gray SW, Lewis N. Internet use leads cancer patients to be active health care consumers. Patient Educ Couns 2010 Dec;81 Suppl:S63-S69. [doi: 10.1016/j.pec.2010.09.004] [Medline: 20889279]

47.   Beckjord EB, Finney Rutten LJ, Squiers L, Arora NK, Volckmann L, Moser RP, et al. Use of the internet to communicate with health care providers in the United States: estimates from the 2003 and 2005 Health Information National Trends Surveys (HINTS). J Med Internet Res 2007;9(3):e20 [FREE Full text] [doi: 10.2196/jmir.9.3.e20] [Medline: 17627929]

48.   Hargittai E. Digital na(t)ives? Variation in Internet skills and uses among members of the "Net Generation". Sociological Inquiry 2010 Jan 21;80(1):92-113. [doi: 10.1111/j.1475-682X.2009.00317.x]

49.   Ramanadhan S, Viswanath K. Health and the information nonseeker: a profile. Health Commun 2006;20(2):131-139. [doi: 10.1207/s15327027hc2002_4] [Medline: 16965250]

50.   Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

51.   Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]

52.   Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]

53.   Lee CJ, Ramírez AS, Lewis N, Gray SW, Hornik RC. Looking beyond the Internet: examining socioeconomic inequalities in cancer information seeking among cancer patients. Health Commun 2012;27(8):806-817. [doi: 10.1080/10410236.2011.647621] [Medline: 22356137]

54.   Lee CJ, Hornik RC. Physician trust moderates the Internet use and physician visit relationship. J Health Commun 2009;14(1):70-76 [FREE Full text] [doi: 10.1080/10810730802592262] [Medline: 19180372]

55.   Zickuhr K. Who's not online and why. Washington, DC: Pew Internet & American Life Project; 2013 Sep 25. URL: http://pewinternet.org/Reports/2013/Non-internet-users/Summary-of-Findings.aspx [accessed 2013-11-20] [WebCite Cache ID 6LH1wWNES]

Abbreviations

HINTS: Health Information National Trends Survey
SES: socioeconomic status
SNS: social networking sites

XSL•FO
RenderX

XSL•FO

RenderX

Original Paper

Online Health Information Seeking Behaviors of Hispanics in New York City: A Community-Based Cross-Sectional Study

Young Ji Lee[1*], RN, MS, PhD; Bernadette Boden-Albala[2*], MPH, Dr PH; Elaine Larson[3,4*], RN, FAAN, PhD; Adam Wilcox[5*], PhD; Suzanne Bakken[3,6*], RN, FAAN, PhD

[1]Division of Health and Biomedical Informatics, Department of Preventive Medicine, Northwestern University, Chicago, IL, United States

[2]Global Institute of Public Health, New York University, New York, NY, United States

[3]School of Nursing, Columbia University, New York, NY, United States

[4]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, United States

[5]Intermountain Healthcare, Murray, UT, United States

[6]Department of Biomedical Informatics, Columbia University, New York, NY, United States

[*]all authors contributed equally

Corresponding Author:
Young Ji Lee, RN, MS, PhD
Division of Health and Biomedical Informatics
Department of Preventive Medicine
Northwestern University
750 N Lake Shore Dr Room 11-186
Chicago, IL, 60611
United States
Phone: 1 312 503 2254
Fax: 1 312 503 5388
Email: young.lee3@northwestern.edu

Abstract

Background: The emergence of the Internet has increased access to health information and can facilitate active individual engagement in health care decision making. Hispanics are the fastest-growing minority group in the United States and are also the most underserved in terms of access to online health information. A growing body of literature has examined correlates of online health information seeking behaviors (HISBs), but few studies have included Hispanics.

Objective: The specific aim of this descriptive, correlational study was to examine factors associated with HISBs of Hispanics.

Methods: The study sample (N=4070) was recruited from five postal zip codes in northern Manhattan for the Washington Heights Inwood Informatics Infrastructure for Comparative Effectiveness Research project. Survey data were collected via interview by bilingual community health workers in a community center, households, and other community settings. Data were analyzed using bivariate analyses and logistic regression.

Results: Among individual respondents, online HISBs were significantly associated with higher education (OR 3.03, 95% CI 2.15-4.29, $P<.001$), worse health status (OR 0.42, 95% CI 0.31-0.57, $P<.001$), and having no hypertension (OR 0.60, 95% CI 0.43-0.84, $P=.003$). Online HISBs of other household members were significantly associated with respondent factors: female gender (OR 1.60, 95% CI 1.22-2.10, $P=.001$), being younger (OR 0.75, 95% CI 0.62-0.90, $P=.002$), being married (OR 1.36, 95% CI 1.09-1.71, $P=.007$), having higher education (OR 1.80, 95% CI 1.404-2.316, $P<.001$), being in worse health (OR 0.59, 95% CI 0.46-0.77, $P<.001$), and having serious health problems increased the odds of their household members' online HISBs (OR 1.83, 95% CI 1.29-2.60, $P=.001$).

Conclusions: This large-scale community survey identified factors associated with online HISBs among Hispanics that merit closer examination. To enhance online HISBs among Hispanics, health care providers and policy makers need to understand the cultural context of the Hispanic population. Results of this study can provide a foundation for the development of informatics-based interventions to improve the health of Hispanics in the United States.

Introduction

Since the Internet has become a promising source of health information for the general public and a target of health information seeking behaviors (HISBs), the use of the Internet for health purposes is an important topic [1-6]. Those living in metropolitan areas with populations larger than 250,000 have been more likely to access the Internet than residents of rural communities [7-9]. Moreover, individuals residing in urban areas are more likely to use the Internet to seek health-related information [3,9]. Although use of the Internet differs by geographic location, social and economic disparities such as geographic distribution of race and ethnicity groups likely account for those differences [9,10].

Hispanics are the most underserved population in the US in terms of access to online health information because of limited opportunities for education and employment and an increased prevalence of poverty [10-13]. According to the US Census Bureau, 16% of the US population (50.5 million) identified themselves as Hispanic or Latino in 2010 [14]. Of Hispanic Internet users, 66% have used the Internet for searching health information. However, non-Hispanic whites and African Americans are more likely to seek health information through the Internet: 73% and 69% respectively [15].

Reducing the inequality of access to health information is linked to reducing and preventing an unequal burden of disease. In 2010, the Department of Health and Human Services launched "Healthy People 2020", which included an objective "to increase the proportion of online health information seekers who report that they can easily access health information" [16]. To meet this objective, the characteristics of individuals who search for health information via the Internet needed to be identified [17,18]. A growing body of literature has examined correlates of online HISBs; however, to our knowledge, there are few such studies about the Hispanic community.

As a part of a larger study, the Washington Heights Inwood Informatics Infrastructure for Comparative Effectiveness Research (WICER) project, our study aimed to examine factors associated with online HISBs among Hispanics.

Methods

Theoretical Framework

Bodie and Dutta's Integrative Model of eHealth Use [19] informed the development of the research question and selection of study variables for our study: What demographic, situational, and literacy factors (health and computer literacy) are associated with online HISBs among Hispanic survey respondents and other members of the same household?

According to the model, variables such as demographics, situational, personal, and cultural factors affect the use of the Internet for obtaining health information [19]. Differences in these variables may contribute to health disparities and a digital divide between people who have and people who do not have access to Internet technology [19]. This model suggests that disparities in social structures such as socioeconomic factors lead to individual-level differences in motivation and online health information seeking ability. The difference in online HISBs causes disparities in lifestyle that are related to health outcomes and continue to contribute to health care disparities [19].

Setting and Sample

The study setting included five zip codes (10031, 10032, 10033, 10034, and 10040) that represent the Washington Heights Inwood community of Northern Manhattan. These communities have been designated as medically underserved areas by the Centers for Medicare and Medicaid Services since they meet the relevant criteria regarding the level of poverty, the proportion of elderly, the incidence of infant mortality, and the ratio of primary care providers to population [20,21]. Currently, 71% of Washington Heights and Inwood area residents are Hispanic [22]. The sample comprised 4070 residents, who completed the WICER household survey between March 2011 and November 2012. Residents who were 18 years or older, English or Spanish speaking, and Hispanic were eligible for inclusion in the study.

Recruitment

After approval by the Columbia University Medical Center Institutional Review Board, recruitment of eligible participants was initiated using multiple methods. Data were collected in households, businesses, or at a designated community space, the Columbia-Community Partnership for Health (CCPH). For the CCPH sample, we recruited a convenience sample of individuals who came to the Center for blood pressure checks or because they were referred by friends. Most of the participants were recruited by a snowball sampling method using respondents' social networks. At the end of the survey interview, interviewers asked participants if they were willing to refer members of their social network for study participation.

Survey Procedures

All interviewers were bilingual and familiar with the Washington Heights and Inwood community. They completed relevant human subjects research training and didactic and field training with the interview guide. Their work was monitored on a daily basis by the survey coordinator who reviewed all interview data. The survey coordinator also conducted regular spot checks of the data collection process in the field. Interviewees were re-trained on an as-needed basis. Before conducting interviews, the trained interviewer obtained informed consent from the participant in their language of choice (English or Spanish). All survey items were self-reported. The interview process took approximately 45 minutes to one hour, and participants received US $25 compensation for their time.

Data Management

All survey data were entered into Lime Survey, a Web-based data management tool, on a secure server. The baseline survey from unique respondents was used in the analyses. Data were

XSL•FO

RenderX

cleaned, and subjects with missing or invalid values for the study variables were removed from the analysis. Data were extracted from Lime Survey into SPSS v. 20.0 for analysis.

Study Variables

Based upon the theoretical framework for the study, the correlates of interest in this study were demographic information, situational factors, health literacy, and computer literacy (Table 1). The two dependent variables in the study were online HISBs of respondents and online HISBs of their household members.

Table 1. Conceptualization and measurement of study variables.

Concept	Variable	Definition	Data type	Instrument
Demographic information	Age, gender, employment, marital status, educational level, insurance		Categorical, Continuous	
Situational factors	Hypertension, health problems, general health status	The specific health situations faced by a patient and their subsequent consumer health information needs	Categorical	Blood pressure question [23], Chronic Burden Scale [24], SF-8 health survey [25]
Health literacy	Health literacy	The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions	Continuous	Newest Vital Sign English or Spanish [26]
Computer literacy	Experiences of social networking	Computer skills and ability to use technology to improve learning, productivity, and performance	Categorical	Use of social networking sites
Online health information seeking behaviors	Online support group or communication with clinicians	The interaction of an individual with or through an electronic device or communication technology to access or transmit health information or to receive guidance and support on a health-related issue	Categorical	Health Information National Trends Survey (HINTS) [27]

Instruments

Demographic and Situational Factors

As summarized in Table 1, demographic and situational data regarding the household respondent and each household member were obtained during the interview. Hypertension was measured by the question, "Have you ever been told by a doctor, nurse, or other health professional that you had hypertension also called high blood pressure or pressure?" [23]. Serious health problems were measured by a question, "Have you experienced any serious personal health problems that have lasted for at least 6 months?" from the Chronic Burden Scale [24]. Self-reported general health status was recorded in five categories: excellent, very good, good, fair, and poor. General health status was measured on a 5-point Likert scale (1=excellent and 5=poor) from the Short Form-8 Health Survey (SF-8) [25].

Health Literacy

The Newest Vital Sign in either English or Spanish (NVS) was used to assess health literacy in the study population. The NVS includes 6 questions to test reading, interpretation, and numeracy skills based on a nutritional label from an ice cream container [26]. This general measure was selected, rather than a more specific measure of eHealth literacy, because of its wide use and WICER's overall goal of understanding the health of the community and social determinants influencing health.

Computer Literacy

The US Department of Education defines computer literacy as "computer skills and ability to use technology to improve learning, productivity, and performance" [28]. However, the definition of computer literacy changes with the technology evolution [29], and several recent studies have demonstrated the positive relationship between use of social networking sites and computer literacy [30,31].

Based on this evidence and the fact that computer literacy was not directly measured in the WICER study, use of social network sites was used as a proxy measure for computer literacy. It was measured by the question, "Do you belong to any social networking sites like Facebook, MySpace, or Twitter?

Online Health Information Seeking Behaviors

Robinson et al (1998) defined interactive health communication as "the interaction of an individual—consumer, patient, caregiver or professional—with or through an electronic device or communication technology to access or transmit health information or to receive guidance and support on a health-related issue" [32]. Based on the definition, this study considered participation in an online support group, email communication with physicians, and using the Internet to look up health or medical information as online HISBs in this study.

Thus, to measure online HISBs, four questions from the Health Information National Trends Survey (HINTS) were used [27]. Respondents' HISBs were measured using three questions: "In

the past 12 months, (1) Have you participated in an online support group for people with similar health or medical issues? (2) Have you used email or the Internet to communicate with a doctor or doctor's office? (3) Have you used the Internet to look up health or medical information?" The HISBs of respondents' household members were measured using a single question: (4) "Does anyone in your household use the Internet to look up health or medical information?" An affirmative response to any of the first three questions was coded as "yes" on online HISBs. An affirmative response to the fourth question was coded as "yes" on household HISBs.

Statistical Analysis

Respondents' online HISBs and those of other household members were analyzed separately. Education was coded as < or ≥ high school, insurance as yes/no, and birthplace as United States or elsewhere. Health literacy scores from NVS were used as a continuous variable [33]. For the NVS, "refused" or "don't know" and missing values were treated as wrong answers and received 0 points.

Initially, univariate analyses were used to examine the frequency and distribution of study variables, calculating mean and standard deviation, range, frequency, and percentage as appropriate. Correlates of respondents' HISBs and those of respondents' household members were analyzed in separate models. Bivariate analyses, including chi-square tests and *t* tests, were used to examine differences between those reporting HISBs and those who did not (separately for respondents and for their household members) in terms of demographic, situational, and literacy (health, computer) variables.

Binary logistic regression analyses were conducted to examine factors associated with online HISBs of respondent and online HISBs of household members. Variables that were significant in the chi-square or *t* test were selected for regression models and entered hierarchically with demographic and situational factors entered first followed by literacy factors. Level of significance for testing of each model was set to an alpha of .05.

Results

Characteristics of Respondents

Demographic Factors

The average age of respondents was 51.2 years old (SD 16.8, range 18-104); 71.33% (2903/4070) were women and 88.80% (3614/4070) were foreign born (Table 2). More than half of the respondents were unemployed (65.18%, 2653/4070), not married (63.10%, 2568/4070), and had a less than high school education (51.74%, 2106/4070). A majority of participants (77.00%, 3134/4070) were Medicare or Medicaid beneficiaries, 15.33% (624/4070) had private or other insurance, and 8.79% (358/4070) were uninsured.

Table 2. Descriptive characteristics of Hispanic participants (N=4070).

Variables	Respondents, n (%)
Demographic factors	
Gender	
Men	1133 (27.84)
Women	2903 (71.33)
Employment status	
Employed	1411 (34.67)
Unemployed	2653 (65.18)
Marital status	
Married/living as	1457 (35.79)
Otherwise	2568 (63.09)
Education	
<High school graduate	2106 (51.74)
≥High school graduate	1906 (46.83)
Birthplace	
Born in the United States	445 (10.93)
Born in other countries	3614 (88.79)
Insurance	
Medicare/Medicaid	3134 (77.00)
Others (veterans, private, etc)	624 (15.33)
None	358 (8.79)
Situational factors	
General health status	
<Good	921 (22.63)
≥Good	3055 (75.06)
Hypertension	
Yes	1608 (39.50)
No	2426 (59.60)
Serious personal health problems	
Yes	342 (8.40)
No	3696 (90.81)
Literacy factors	
Social networking sites	
Yes	867 (21.30)
No	2815 (69.16)
Online health information seeking behaviors	
Respondent	317 (7.89)
Household member	466 (11.45)
Demographic factors	
Age, mean (SD)	51.2 (16.81)
Literacy factors	
Health literacy, mean (SD)	2.1 (1.96)

Situational Factors

Most respondents reported their general health status as at least "good" (75.06%, 3055/4070) and without serious health problems (90.81%, 3696/4070). A large proportion of respondents (39.51%, 1608/4070) answered that they had been diagnosed with hypertension by a clinician.

Literacy Factors

The mean NVS score was 2.2 (SD 1.96), indicating the possibility of marginal or inadequate literacy. For computer literacy, only 21.30% of respondents (867/4070) answered that they had used social networking sites.

Online Health Information Seeking Behaviors

Only 7.79% of respondents (317/4070) reported at least one of three HISBs (ie, online support group, email communication with physician, used the Internet to search for health-related information); 11.45% (466/4070) reported that a household member had used the Internet to search for health-related information.

Factors Associated With Online Health Information Seeking Behaviors

Primary Respondent

Several demographic and situational factors were independently associated with respondent online HISBs, including having higher education (OR 3.03, 95% CI 2.15-4.29, $P<.001$), being in worse health status (OR 0.42, 95% CI 0.31-0.57, $P<.001$), and having no hypertension (OR 0.60, 95% CI 0.43-0.84, $P=.003$). Social networking site users were more than three times more likely than non-users to seek health information online (OR 3.78, 95% CI 2.78- 5.13, $P<.001$) (Table 3). Model fit was poor (Hosmer and Lemeshow χ^2_8=17.78, $P=.02$).

Table 3. Binary logistic regression: correlates of respondents' online HISBs (N=4070).

Respondent factors	Model OR (95% CI)	P value
Demographic and situational factors		
Gender	1.33 (0.96-1.83)	.08
Age	1.14 (0.92-1.41)	.25
General health status	0.42 (0.31-0.57)[a]	<.001
Hypertension	0.60 (0.43-0.84)[b]	.003
Insurance	0.81 (0.54-1.22)	.31
Employment status	1.17 (0.88-1.54)	.29
Education level	3.03 (2.15-4.29)[a]	<.001
Nativity	1.25 (0.88-1.77)	.22
Literacy factors		
Health literacy level (NVS)	0.99 (0.93-1.07)	.87
Use of SNS	3.78 (2.78-5.13)[a]	<.001
Hosmer and Lemeshow χ^2_8	17.78[c]	.02

[a]$P<.001$.

[b]$P<.01$.

[c]$P<.05$.

Other Household Members

Respondents' use of social networking was a significant factor (OR 2.24, 95% CI 1.74-2.89, $P<.001$), controlling for other factors, in predicting household members HISBs. Additionally, several respondents' characteristics were associated with their household members online HISBs: female (OR 1.60, 95% CI 1.22-2.10, $P=.001$), younger (OR 0.75, 95% CI 0.62-0.90, $P=.002$), married (OR 1.36, 95% CI 1.09-1.71, $P=.007$), higher education (OR 1.80, 95% CI 1.404-2.316, $P<.001$), being in worse health status (OR 0.59, 95% CI 0.46-0.77, $P<.001$), and having serious health problem (OR 1.83, 95% CI 1.29-2.60, $P=.001$) (Table 4). The model for household members' online HISBs demonstrated good fit (Hosmer and Lemeshow χ^2_8=6.31, $P=.66$).

Table 4. Binary logistic regression: correlates of household members' online HISBs (N=4070).

Respondent factors	Model OR (95% CI)	P value
Demographic and situational factors		
Gender	1.60 (1.22-2.10)[a]	.001
Age	0.75 (0.62-0.90)[a]	.002
General health status	0.59 (0.46-0.77)[b]	<.001
Hypertension	0.93 (0.72-1.21)	.60
Marital status	1.36 (1.09-1.71)[a]	.007
Employment status	1.18 (0.94-1.49)	.16
Education level	1.80 (1.40-2.32)[b]	<.001
Serious health problem	1.83 (1.29-2.60)[a]	.001
Literacy factors		
Health literacy level (NVS)	0.97 (0.91-1.03)	.26
Use of SNS	2.24 (1.74-2.90)[b]	<.001
Hosmer and Lemeshow χ^2_8	5.85	.66

[a]$P<.01$.

[b]$P<.001$.

Discussion

Principal Results

Summary

Most Hispanics have sought health information from family and friends or community groups in the past year [34]. Recently, there is increased reliance on the Internet. The National Health Interview Survey conducted by National Center for Health Statistics showed that 28.8% of Hispanics aged 18-64 used the Internet to find health-related information [35]. However, only 7.79% of our survey respondents reported online HISBs, which is significantly lower than the national data [35]. There are several potential reasons for this difference. First, Washington Height and Inwood are designated as medically underserved areas. Another study conducted in a medically underserved area showed that only 21% of respondents accessed to the Internet for health information [21]. This suggests that there may be a significant disparity in online HISBs between underserved populations and the general population [21]. Second, there may also be inequalities in access to online health information within Hispanics due to the heterogeneity of that population [11,26,36,37], which may explain differences between our findings and the national data.

This study provides new data regarding correlates of the use of the Internet by Hispanics for seeking health information. Findings are summarized and then discussed according to the concepts from the Integrative Model of eHealth Use [19].

Higher education level, being in worse (poor/fair) health status, having normal blood pressure (ie, no hypertension), and being computer literate were positively associated with online HISBs. However, the final model of respondents' online HISBs demonstrated poor fit. Given the relatively large sample size of this study (N=4070), small differences across the sample can influence model fit [38]. Thus, this is one potential reason for the poor model fit. Other potential reasons for the poor model fit are missing variables that are associated with online HISBs and operationalization of the study measures.

Furthermore, respondents' demographic, situational, and computer literacy factors were associated with increased odds of their household members' online HISBs. Household members were more likely to seek health-related information through the Internet when the respondents were female, younger, married, highly educated, computer literate, in worse health status, and had serious health problems. In contrast to the respondent model, this model demonstrated good fit with the data.

Demographic Characteristics

Studies of the general population in the United States have shown that being female [4,39-43], being younger [4,7,40,43-47], and having more education are positively associated with online HISBs [3,4,7,18,39-41,44-49]. Studies on Hispanics have shown similar results [11,50]. Our findings are consistent with previous studies that showed that better-educated respondents were more likely to access health information through the Internet. In particular, Miller et al found that Hispanics had the strongest relationship between education and online HISBs among ethnic and racial groups [51].

However, some of the findings were inconsistent with those reported in the literature. Age did not influence respondents' online HISBs in our study, whereas earlier studies showed that younger age is positively related to online HISBs [4,7,40,43-47,50]. This may be because the great majority of the respondents were over 40 years old. Among the survey respondents, 20.74 % were over 65 years old (844/4070), and 51.33% of respondents were between 40 and 64 years old

(2089/4070). Recently, the older adults who use the Internet as a source of health information have been increasing. About 69% of the population over age 65 report online HISBs [52,53]. The number will continue to increase, since the majority of online health information seekers are adults between 40 and 59 years old [53,54]. In addition, the US census showed that among Internet users, older people are more engaged in online HISBs than younger people [55].

In contrast to our findings, previous studies showed that females were more likely to search health information through the Internet [4,11,39-43]. One study found that men and people without children were more likely to seek health information for themselves rather than others [56]. The relatively small proportion of males (27.84%, 1133/4070) coupled with the low online HISB may have influenced the ability to detect gender differences even in our large sample. However, our findings suggest that respondents' gender and marital status are associated with household members' online HISBs. Women play a key role in managing the health of the Hispanic family [11]; for example, household members may seek health information at a female family member's request. The relationship between respondents' marital status and their household members' online HISBs is consistent with Sadasivam's finding that being married was positively associated with surrogate-seekers' online HISBs [57]. In that study, a surrogate seeker was defined as a person who looked for health information for family members or friends [57].

Respondents' higher education level was positively associated with household members' online HISBs. Recent studies have shown that children of less-educated parents are less likely to seek health information [58]. In our study, about one third of household members who had sought health-related information were sons, daughters, or grandchildren 33.9% (158/466) of the respondent.

Situational Factors

Our study showed that individuals with poor health status were more likely to seek health information through the Internet. This is consistent with previous studies that have found that people with poor health status may have stronger needs for information [59,60]. Online health information can meet their higher demand for health information because of easy access. Their needs for health information may lead them to use online health information to manage their health [60].

Respondents without hypertension (ie, normal blood pressure) were more likely to seek health information through the Internet. Previous studies have found that having chronic disease including hypertension was positively associated with online HISBs [48,61]. However, Ayers and Kronenfeld suggested that online HISBs are not merely affected by the presence of a particular chronic illness, but rather by the total number of chronic conditions [62]. Furthermore, several studies have found that individuals who have hypertension seek health information less than those with other chronic diseases [54]. A Pew Internet survey showed that among the online health information seekers with one or more chronic conditions, the percentage of hypertension patients (57%) was less than that of cancer patients (62%) or lung patients (68%) [63]. These findings provide a

possible rationale for the association between hypertension status and online HISBs.

For the household members, it is not possible to determine from this study if the household member looked up health information for themselves, the respondents, or another friend or family member. Household members may have sought health information for respondents since respondents suffered from serious health problems and they perceived their health status as poor. Familism, an important Hispanic cultural value with implications for the engagement of family members in the care of a patient, is a possible rationale for this finding [11,36]. Furthermore, most survey respondents were immigrants. Among immigrants, the family plays an important role in HISBs. Instead of consulting with health care providers, they often ask their family members about health information and for advice [64,65].

Literacy Factors

An individual who is computer literate is more likely to go to the Internet for finding health information. Several studies have shown that the ability to use a computer is related to online HISBs [66]. In our study, computer literacy as measured by social networking was positively associated with online HISBs. Moreover, respondents' computer literacy was also positively associated with household members' online HISBs. This is consistent with a study that found that parents' computer literacy may affect children's computer use [67]. Respondents in our study answered that 33.9% of their household members who went to the Internet for health information were their children.

Health literacy was not significantly associated with online HISBs of respondents or their household members, although it has consistently been identified in the literature as a challenge when people use the Internet to search for health information [68-71]. A possible explanation for the lack of significance is the floor effect [33] because most respondents scored low on the NVS.

Limitations

There are several limitations to this study. Generalizability is a potential limitation of this study due to the non-probability sampling method. Although the study adopted several sampling methods for recruitment, most participants were recruited using non-probability sampling, and the resulting sample is more Hispanic, female, and older than the Washington Heights and Inwood population. Moreover, this study recruited Hispanics who lived in urban areas in New York City; therefore, the findings may not be generalizable to Hispanics living in rural areas or in other cities.

Second, because our study relied on self-reported information, social desirability is a potential concern. Sometimes respondents tend to answer in a way that they think the researcher wants. They tend to over-report for the desired behaviors such as physical activity and under- report undesirable ones such as alcohol consumption [72]. Furthermore, the question regarding household members' online HISB was answered by respondents on behalf of household members. Therefore, it is possible that the percentage of online HISBs among household members may not be accurate.

Third, this study did not explicitly identify respondents who accessed the Internet from their cell phones. A Pew Hispanic Center reported that Hispanics are more likely than non-Hispanics to access the Internet through mobile devices; 76% of Hispanics access the mobile Internet compared to 60% of non-Hispanics [73]. Therefore, this study may have underestimated online HISBs among the survey respondents. An alternative explanation is that our use of bilingual data collectors resulted in a sample that included individuals not typically included in other studies.

Implications

Online HISBs can lead patients to make informed health care decisions by increasing their participation in health management. Those decisions may affect the relationship with health care providers [62]. To assist patients, there is a need for a health information infrastructure for shared decision making between patients and the health care system [74]. However, decision-making interventions may not be successfully implemented if discrepancies exist across populations [75]. Our study of online HISBs among Hispanics may suggest what needs to be considered to resolve the discrepancy and to implement an infrastructure in Hispanic community such as the one that we studied.

However, we showed that overall access for online health information was lower than national data. Government agencies and policy makers need to understand unique characteristics of Hispanic communities to design strategies and interventions for equitable access to online health information among underserved populations. This understanding may lead governments to develop policies to allocate and disseminate infrastructures and resources [74,76]. Large-scale improvement in the Internet technology infrastructure will reduce costs and barriers to accessing health information [77]. To provide timely and accurate health information across populations, policy makers need to consider a broad spectrum of health literacy levels and cultural issues of the designated population [78].

Conclusions

This is the first large-scale study of online HISBs in the Hispanic population. This study not only confirmed factors associated with online HISBs identified in the literature but also revealed findings that were not previously described. Also, this study showed that respondents' demographic, situational, and literacy factors were significantly associated with their household members' online HISBs. Studies of online HISBs need to consider various associated factors to explain behavior [79]. To increase the number of online health information seekers among Hispanics, cultural values such as familism need to be considered. Results of this study can provide a foundation for the development of informatics-based interventions to improve the health of Hispanics in the United States.

Conflicts of Interest

None declared.

References

1. Alpay L, Verhoef J, Toussaint P, Zwetsloot-Schonk B. What makes an "informed patient"? The impact of contextualization on the search for health information on the Internet. Stud Health Technol Inform 2006;124:913-919. [Medline: 17108628]
2. Anker AE, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. Patient Educ Couns 2011 Mar;82(3):346-354. [doi: 10.1016/j.pec.2010.12.008] [Medline: 21239134]
3. Flynn KE, Smith MA, Freese J. When do older adults turn to the internet for health information? Findings from the Wisconsin Longitudinal Study. J Gen Intern Med 2006 Dec;21(12):1295-1301 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00622.x] [Medline: 16995892]
4. Koch-Weser S, Bradshaw YS, Gualtieri L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. J Health Commun 2010;15 Suppl 3:279-293. [doi: 10.1080/10810730.2010.522700] [Medline: 21154099]
5. Nauert RF. Telehealth Consumer-Provider Interaction: A Chronic Disease Intervention in an Underserved Population. Austin, TX: The University of Texas at Austin; 2002.
6. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396 [FREE Full text] [doi: 10.1093/her/cyl096] [Medline: 16971674]
7. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]
8. Ruggiero KJ, Gros DF, McCauley J, de Arellano MA, Danielson CK. Rural adults' use of health-related information online: data from a 2006 National Online Health Survey. Telemed J E Health 2011 Jun;17(5):329-334 [FREE Full text] [doi: 10.1089/tmj.2010.0195] [Medline: 21524201]

9.   Hale TM, Cotten SR, Drentea P, Goldner M. Rural-urban differences in general and health-related Internet use. American Behavioral Scientist 2010 Mar 09;53(9):1304-1325. [doi: 10.1177/0002764210361685]

10. White K, Haas JS, Williams DR. Elucidating the role of place in health care disparities: the example of racial/ethnic residential segregation. Health Serv Res 2012 Jun;47(3 Pt 2):1278-1299 [FREE Full text] [doi: 10.1111/j.1475-6773.2012.01410.x] [Medline: 22515933]

11. Peña-Purcell N. Hispanics' use of Internet health information: an exploratory study. J Med Libr Assoc 2008 Apr;96(2):101-107 [FREE Full text] [doi: 10.3163/1536-5050.96.2.101] [Medline: 18379664]

12. Ginossar T, Nelson S. Reducing the health and digital divides: A model for using community-based participatory research approach to e-Health interventions in low-income Hispanic communities. J Comput Mediat Commun 2010;15(4):530-551. [doi: 10.1111/j.1083-6101.2009.01513.x]

13. 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]

14. The U.S. Census Bureau. The Hispanic Population. 2010. URL: http://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf [accessed 2014-04-28] [WebCite Cache ID 6PA3ulLIl]

15. Fox SD, Duggan M. Health Online. 2013. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2014-04-28] [WebCite Cache ID 6PA4GVwtT]

16. The United States Department of Health and Human Service (USDHHS). Healthy People 2020. URL: http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=18 [accessed 2014-04-28] [WebCite Cache ID 6PA4MsGA7]

17. Kalichman SC, Cherry C, Cain D, Pope H, Kalichman M, Eaton L, et al. Internet-based health information consumer skills intervention for people living with HIV/AIDS. J Consult Clin Psychol 2006 Jun;74(3):545-554. [doi: 10.1037/0022-006X.74.3.545] [Medline: 16822111]

18. McInnes DK, Gifford AL, Kazis LE, Wagner TH. Disparities in health-related internet use by US veterans: results from a national survey. Inform Prim Care 2010;18(1):59-68. [Medline: 20429979]

19. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: eHealth literacy, health disparities, and the digital divide. Health Mark Q 2008;25(1-2):175-203. [doi: 10.1080/07359680802126301] [Medline: 18935884]

20. Pati S, Romero D, Chavkin W. Changes in use of health insurance and food assistance programs in medically underserved communities in the era of welfare reform: an urban study. Am J Public Health 2002 Sep;92(9):1441-1445. [Medline: 12197970]

21. Zach L, Dalrymple PW, Rogers ML, Williver-Farr H. Assessing internet access and use in a medically underserved population: implications for providing enhanced health information services. Health Info Libr J 2012 Mar;29(1):61-71. [doi: 10.1111/j.1471-1842.2011.00971.x] [Medline: 22335290]

22. New York City Department of Health and Mental Hygiene (NYCDHMH). Community Health Profiles: New York City Department of Health and Mental Hygiene. 2006. URL: http://www.nyc.gov/html/doh/downloads/pdf/data/2006chp-301.pdf [accessed 2014-04-28] [WebCite Cache ID 6PA4unwPc]

23. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Questionnaire. 2012. URL: http://www.cdc.gov/brfss/questionnaires/pdf-ques/2013%20BRFSS_English.pdf [accessed 2014-04-28] [WebCite Cache ID 6PA4yBf62]

24. Mujahid MS, Diez Roux AV, Cooper RC, Shea S, Williams DR. Neighborhood stressors and race/ethnic differences in hypertension prevalence (the Multi-Ethnic Study of Atherosclerosis). Am J Hypertens 2011 Feb;24(2):187-193 [FREE Full text] [doi: 10.1038/ajh.2010.200] [Medline: 20847728]

25. Yen ST, Chen Z, Eastwood DB. Lifestyles, demographics, dietary behavior, and obesity: a switching regression analysis. Health Serv Res 2009 Aug;44(4):1345-1369 [FREE Full text] [doi: 10.1111/j.1475-6773.2009.00969.x] [Medline: 19490158]

26. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]

27. Health Information National Trends Survey (HINTS). SurveyInstruments. 2007. URL: http://hints.cancer.gov/instrument.aspx [accessed 2014-04-28] [WebCite Cache ID 6PA51snsO]

28. Bers MU. Beyond computer literacy: supporting youth's positive development through technology. New Dir Youth Dev 2010;2010(128):13-23. [doi: 10.1002/yd.371] [Medline: 21240949]

29. Lin TC. A computer literacy scale for newly enrolled nursing college students: development and validation. J Nurs Res 2011 Dec;19(4):305-317. [doi: 10.1097/JNR.0b013e318236d03f] [Medline: 22089656]

30. Appel M. Are heavy users of computer games and social media more computer literate? Computers & Education 2012 Dec;59(4):1339-1349. [doi: 10.1016/j.compedu.2012.06.004]

31. Smith W, Bedayse S, Lalwah SL, Paryag A. Computer literacy and attitudes of dental students and staff at the University of the West Indies Dental School. Eur J Dent Educ 2009 Aug;13(3):179-183. [doi: 10.1111/j.1600-0579.2009.00571.x] [Medline: 19630937]

32. Robinson TN, Patrick K, Eng TR, Gustafson D. An evidence-based approach to interactive health communication: a challenge to medicine in the information age. Science Panel on Interactive Communication and Health. JAMA 1998 Oct 14;280(14):1264-1269. [Medline: 9786378]

XSL•FO
RenderX

33. Everitt BS. The Cambridge dictionary of statistics. Cambridge: Cambridge University Press; 2010.

34. Livingston G, Minushkin S, Cohn D. Hispanics and health care in the United States: access, information and knowledge. 2008. URL: http://www.pewhispanic.org/2008/08/13/hispanics-and-health-care-in-the-united-states-access-information-and-knowledge/ [accessed 2014-05-23] [WebCite Cache ID 6PmBnuxMc]

35. Cohen RA, Adams PF. Use of the Internet for Health Information: United States. 2009. URL: http://www.cdc.gov/nchs/data/databriefs/db66.pdf [accessed 2014-04-28] [WebCite Cache ID 6PA55jSDu]

36. Cheong PH. Health communication resources for uninsured and insured Hispanics. Health Commun 2007;21(2):153-163. [doi: 10.1080/10410230701307188] [Medline: 17523861]

37. Minor M. Commentary – Comparing the Hispanic and Non-Hispanic markets: How different are they? Journal of Services Marketing 1992;6(2):29-32. [doi: 10.1108/08876049210035836]

38. University of Strathclyde. Goodness of Fit Measures. 2013. URL: http://www.strath.ac.uk/aer/materials/5furtherquantitativeresearchdesignandanalysis/unit6/goodnessoffitmeasures/ [accessed 2014-04-28] [WebCite Cache ID 6PA59QViY]

39. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]

40. Gallagher S, Doherty DT. Searching for health information online: characteristics of online health seekers. J Evid Based Med 2009 May;2(2):99-106. [doi: 10.1111/j.1756-5391.2009.01021.x] [Medline: 21348996]

41. Jensen JD, King AJ, Davis LA, Guntzviller LM. Utilization of internet technology by low-income adults: the role of health literacy, health numeracy, and computer assistance. J Aging Health 2010 Sep;22(6):804-826. [doi: 10.1177/0898264310366161] [Medline: 20495159]

42. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. Int J Med Inform 2006 Jan;75(1):8-28. [doi: 10.1016/j.ijmedinf.2005.07.032] [Medline: 16125453]

43. Renahy E, Parizot I, Chauvin P. Health information seeking on the Internet: a double divide? Results from a representative survey in the Paris metropolitan area, France, 2005-2006. BMC Public Health 2008;8:69 [FREE Full text] [doi: 10.1186/1471-2458-8-69] [Medline: 18291024]

44. Chou WY, Liu B, Post S, Hesse B. Health-related Internet use among cancer survivors: data from the Health Information National Trends Survey, 2003-2008. J Cancer Surviv 2011 Sep;5(3):263-270. [doi: 10.1007/s11764-011-0179-5] [Medline: 21505861]

45. Cohall AT, Nye A, Moon-Howard J, Kukafka R, Dye B, Vaughan RD, et al. Computer use, internet access, and online health searching among Harlem adults. Am J Health Promot 2011;25(5):325-333. [doi: 10.4278/ajhp.090325-QUAN-121] [Medline: 21534835]

46. 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;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]

47. van Uden-Kraan CF, Drossaert CH, Taal E, Smit WM, Moens HJ, Siesling S, et al. Health-related Internet use by patients with somatic diseases: frequency of use and characteristics of users. Inform Health Soc Care 2009 Jan;34(1):18-29. [doi: 10.1080/17538150902773272] [Medline: 19306196]

48. Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? Health Serv Res 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00510.x] [Medline: 16704514]

49. Reinfeld-Kirkman N, Kalucy E, Roeger L. The relationship between self-reported health status and the increasing likelihood of South Australians seeking Internet health information. Aust N Z J Public Health 2010 Aug;34(4):422-426. [doi: 10.1111/j.1753-6405.2010.00576.x] [Medline: 20649784]

50. Jesus MD, Xiao C. Predicting Internet use as a source of health information: a "Language Divide" among the Hispanic population in the United States. Policy & Internet 2012;4(2):1-11. [doi: 10.1515/1944-2866.1178]

51. Miller EA, West DM, Wasserman M. Health information Websites: characteristics of US users by race and ethnicity. J Telemed Telecare 2007;13(6):298-302. [doi: 10.1258/135763307781644915] [Medline: 17785031]

52. Fox S. Health topics. 2011. URL: http://www.pewinternet.org/2011/02/01/health-topics-2/ [accessed 2014-04-28] [WebCite Cache ID 6PA5D4ddK]

53. Huang M. Older adults' online health information seeking behavior. 2012 Presented at: iConference 2012; February 7-10, 2012; Toronto, Canada. [doi: 10.1145/2132176.2132220]

54. Gray K, Elliott K, Wale J. A community education initiative to improve using online health information: participation and impact. Inform Health Soc Care 2013 Sep;38(3):171-181. [doi: 10.3109/17538157.2012.705201] [Medline: 23324099]

55. United States Census Bureau. Computer and Internet Use in the United States. 2013. URL: https://www.census.gov/prod/2013pubs/p20-569.pdf [accessed 2014-04-28] [WebCite Cache ID 6PA5JFFyS]

56. Stern MJ, Cotten SR, Drentea P. The separate spheres of online health: Gender, parenting, and online health information searching in the information age. Journal of Family Issues 2011 Oct 21;33(10):1324-1350. [doi: 10.1177/0192513X11425459]

57. Sadasivam RS, Kinney RL, Lemon SC, Shimada SL, Allison JJ, Houston TK. Internet health information seeking is a team sport: analysis of the Pew Internet Survey. Int J Med Inform 2013 Mar;82(3):193-200. [doi: 10.1016/j.ijmedinf.2012.09.008] [Medline: 23149121]

58. 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]

59. Shim JW. The use of e-Health information and health behavior change: The role of perceived health status and types of e-Health information use. TOCOMMJ 2008 Oct 30;2(1):156-163. [doi: 10.2174/1874916X00802010156]

60. Xiao N, Sharman R, Rao H, Upadhyaya S. Factors influencing online health information search: An empirical analysis of a national cancer-related survey. Decision Support Systems 2014 Jan;57:417-427. [doi: 10.1016/j.dss.2012.10.047]

61. Wagner TH, Baker LC, Bundorf MK, Singer S. Use of the Internet for health information by the chronically ill. Prev Chronic Dis 2004 Oct;1(4):A13 [FREE Full text] [Medline: 15670445]

62. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]

63. Fox SP, Purcell K. Chronic disease and the Internet. 2010. URL: http://www.pewinternet.org/2010/03/24/chronic-disease-and-the-internet/ [accessed 2014-04-28] [WebCite Cache ID 6PA5MXprQ]

64. Leclere FB, Jensen L, Biddlecom AE. Health care utilization, family context, and adaptation among immigrants to the United States. J Health Soc Behav 1994 Dec;35(4):370-384. [Medline: 7844331]

65. Livingston G. The Latino digital divide: The native born versus the foreign born. 2010. URL: http://www.pewhispanic.org/files/reports/123.pdf [accessed 2014-04-28] [WebCite Cache ID 6PA5PhrsO]

66. Kahn JS, Aulakh V, Bosworth A. What it takes: characteristics of the ideal personal health record. Health Aff (Millwood) 2009;28(2):369-376 [FREE Full text] [doi: 10.1377/hlthaff.28.2.369] [Medline: 19275992]

67. Dincer S. A study of the relationship between pupils and parents' computer literacy level and use. Procedia - Social and Behavioral Sciences 2012 Jan;46:484-489. [doi: 10.1016/j.sbspro.2012.05.146]

68. Benigeri M. Shortcomings of health information on the Internet. Health Promotion International 2003 Dec 01;18(4):381-386. [doi: 10.1093/heapro/dag409]

69. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]

70. 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]

71. Pluye P, Grad R, Repchinsky C, Jovaisas B, Johnson-Lafleur J, Carrier M, et al. Four levels of outcomes of information-seeking: A mixed methods study in primary health care. J Am Soc Inf Sci Tec 2012 Dec 10;64(1):108-125. [doi: 10.1002/asi.22793]

72. Donaldson SI, Grant-vallone EJ. Understanding self-report bias in organizational behavior research. JBus Psychol 2002;17:245-260. [doi: 10.1023/A:1019637632584]

73. Lopez MH, Gonzalez-Barrera A, Patten E. Closing the digital divide: Latinos and technology adoption. 2013. URL: http://www.pewhispanic.org/files/2013/03/Latinos_Social_Media_and_Mobile_Tech_03-2013_final.pdf [accessed 2014-04-28] [WebCite Cache ID 6PAGjQy63]

74. Lorence DP, Park H, Fox S. Racial disparities in health information access: resilience of the Digital Divide. J Med Syst 2006 Aug;30(4):241-249. [Medline: 16978003]

75. Lorence D, Park H. Group disparities and health information: a study of online access for the underserved. Health Informatics J 2008 Mar;14(1):29-38. [doi: 10.1177/1460458207086332] [Medline: 18258673]

76. Xiao N, Sharman R, Rao H, Upadhyaya S. Factors influencing online health information search: An empirical analysis of a national cancer-related survey. Decision Support Systems 2014 Jan;57:417-427. [doi: 10.1016/j.dss.2012.10.047]

77. Chou WY, Liu B, Post S, Hesse B. Health-related Internet use among cancer survivors: data from the Health Information National Trends Survey, 2003-2008. J Cancer Surviv 2011 Sep;5(3):263-270. [doi: 10.1007/s11764-011-0179-5] [Medline: 21505861]

78. Tu HT, Cohen GR. Striking jump in consumers seeking health care information. 2008. URL: http://hschange.org/CONTENT/1006/ [accessed 2014-05-23] [WebCite Cache ID 6PmCdUeWp]

79. Caiata-Zufferey M, Abraham A, Sommerhalder K, Schulz PJ. Online health information seeking in the context of the medical consultation in Switzerland. Qual Health Res 2010 Aug;20(8):1050-1061. [doi: 10.1177/1049732310368404] [Medline: 20442347]

Abbreviations

CCPH: Columbia-Community Partnership for Health
HINTS: Health Information National Trends Survey
HISB: health information seeking behaviors
NVS: Newest Vital Sign
SF-8: Short Form-8 survey

WICER: Washington Heights Inwood Informatics Infrastructure for Comparative Effectiveness Research

XSL·FO
RenderX

<u>Original Paper</u>

A Web-Based Non-Intrusive Ambient System to Measure and Classify Activities of Daily Living

Reto A Stucki[1], MSc; Prabitha Urwyler[1], PhD; Luca Rampa[2], MD; René Müri[1,3], MD; Urs P Mosimann[1,2], MD, PhD; Tobias Nef[4,5], PhD

[1]Gerontechnology and Rehabilitation Group, University of Bern, Bern, Switzerland

[2]University Hospital of Old Age Psychiatry, University of Bern, Bern, Switzerland

[3]Perception and Eye Movement Laboratory, Departments of Neurology and Clinical Research, University Hospital Inselspital, Bern, Switzerland

[4]Gerontechnology and Rehabilitation Group, ARTORG Center for Biomedical Engineering Research, Bern, Switzerland

[5]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

Corresponding Author:
Tobias Nef, PhD
Gerontechnology and Rehabilitation Group
ARTORG Center for Biomedical Engineering Research
University of Bern
Bern, 3010
Switzerland
Phone: 41 31 632 46 97
Fax: 41 31 632 75 76
Email: tobias.nef@artorg.unibe.ch

Abstract

Background: The number of older adults in the global population is increasing. This demographic shift leads to an increasing prevalence of age-associated disorders, such as Alzheimer's disease and other types of dementia. With the progression of the disease, the risk for institutional care increases, which contrasts with the desire of most patients to stay in their home environment. Despite doctors' and caregivers' awareness of the patient's cognitive status, they are often uncertain about its consequences on activities of daily living (ADL). To provide effective care, they need to know how patients cope with ADL, in particular, the estimation of risks associated with the cognitive decline. The occurrence, performance, and duration of different ADL are important indicators of functional ability. The patient's ability to cope with these activities is traditionally assessed with questionnaires, which has disadvantages (eg, lack of reliability and sensitivity). Several groups have proposed sensor-based systems to recognize and quantify these activities in the patient's home. Combined with Web technology, these systems can inform caregivers about their patients in real-time (eg, via smartphone).

Objective: We hypothesize that a non-intrusive system, which does not use body-mounted sensors, video-based imaging, and microphone recordings would be better suited for use in dementia patients. Since it does not require patient's attention and compliance, such a system might be well accepted by patients. We present a passive, Web-based, non-intrusive, assistive technology system that recognizes and classifies ADL.

Methods: The components of this novel assistive technology system were wireless sensors distributed in every room of the participant's home and a central computer unit (CCU). The environmental data were acquired for 20 days (per participant) and then stored and processed on the CCU. In consultation with medical experts, eight ADL were classified.

Results: In this study, 10 healthy participants (6 women, 4 men; mean age 48.8 years; SD 20.0 years; age range 28-79 years) were included. For explorative purposes, one female Alzheimer patient (Montreal Cognitive Assessment score=23, Timed Up and Go=19.8 seconds, Trail Making Test A=84.3 seconds, Trail Making Test B=146 seconds) was measured in parallel with the healthy subjects. In total, 1317 ADL were performed by the participants, 1211 ADL were classified correctly, and 106 ADL were missed. This led to an overall sensitivity of 91.27% and a specificity of 92.52%. Each subject performed an average of 134.8 ADL (SD 75).

Conclusions: The non-intrusive wireless sensor system can acquire environmental data essential for the classification of activities of daily living. By analyzing retrieved data, it is possible to distinguish and assign data patterns to subjects' specific activities and

XSL•FO
RenderX

to identify eight different activities in daily living. The Web-based technology allows the system to improve care and provides valuable information about the patient in real-time.

Introduction

The number of older adults in the global population is increasing both in absolute and relative terms due to increased life expectancy [1]. This demographic shift has led to an increasing prevalence of age-associated disorders, such as Alzheimer's disease and other types of dementia. According to the World Health Organization, today there are 35 million patients worldwide suffering from dementia. By the year 2030, the number of patients is expected to double to 66 million worldwide [2]. Alzheimer's disease is a neurodegenerative disease, and with its progression, the ability to cope with activities of daily living decreases, leading to reduced autonomy and increased need for care [3]. With the progression of the disease, the risk of institutional care increases, which contrasts with the desire of most patients who would like to live autonomously in their familiar home environment as long as possible [4]. Despite awareness of the patient's cognitive status by doctors and caregivers, they are often uncertain about its consequences on activities of daily living (ADL). To provide effective care, they need to know how good patients cope with ADL, in particular when it comes to the estimation of risks associated with the cognitive decline. It is not yet clear how cognitive decline influences ADL. In this regard, the occurrence, performance, and duration of different ADL are important indicators of functional ability [3].

The patient's ability to cope with ADL is traditionally assessed with questionnaires (eg, Stanford Health Assessment Questionnaire [5] or the Barthel ADL (activities of daily living) Index [6]), and it is important information for professional caregivers in order to optimize medication and personalize care [3]. However, using questionnaires to assess behavioral data in patients with cognitive impairments has disadvantages (eg, lack of reliability and sensitivity), and several groups have proposed sensor-based systems to recognize and quantify ADL [4,7-9] in the patient's home. Besides storing the occurrence, duration, and location of ADL, when combined with Web technology, these systems can also inform caregivers in real-time (eg, via smartphone) about the current ADL of their patients [7-10]. Moreover, current and future clinical trials of new drug interventions, in Alzheimer's disease will need to prove their effects on ADL function, and sensitive and reliable measurements will be of great importance. Also, it has been suggested that such systems could facilitate and support aging-in-place and improve medical care [9,10].

One possible technical solution for recognizing ADL is using comprehensive sensor networks to measure patient's activities. In so-called smart homes, several sensors, for example, accelerometers, microphone arrays, pressure sensitive mats, gas sensors, and cameras are installed in the proximity of older patients to determine specific activities and to monitor their ability to cope with ADL [7,8]. Commercial smart-home monitoring systems for dementia patients are already available on the market. For example, ALARM-NET and CareWatch have been developed to recognize ADL and to improve dementia care provided by both formal and informal caregivers [9,10]. Other technical systems make use of detachable sensor arrays, for example, wearable accelerometers, infrared sensors and pressure sensors in furniture, wireless heart rate monitors, or radio frequency technology. Such sensor networks are used for automated ADL classification in dementia patients to maintain their autonomy, for recognizing emergencies, and monitoring the progression of their disease [11-16].

However, state-of-the-art technology systems have some limitations. Smart homes, for example, are often very comprehensive and require significant installation efforts, which makes these systems costly and challenging to install in older houses [17]. Also, most smart home systems use video-based technology or microphone recordings, which might conflict with the patient's privacy [18]. Other approaches require that the patient wear body-mounted sensors or interact with a system for data acquisition [11,12,15]. Using these systems requires patient compliance, which is not always assured when patients with cognitive impairments are overloaded by technology [19,20]. Furthermore, others have reported that too-intrusive systems are not well accepted by patients or caregivers [21,22].

We hypothesize that a non-intrusive system, which does not use body-mounted sensors, avoids video-based imaging and microphone recordings, and does not require any interaction or patient compliance would be better suited for use in dementia patients. If small, easy to install, and not in need of patient's attention and compliance, such a system might be well accepted by patients [23]. Furthermore, such a system, due to its low cost, robustness, and easy-to-use approach, can be of interest for other researchers from different fields. ADL can be a relevant outcome measurement for future drug intervention studies.

In this paper, we present a novel passive, Web-based, non-intrusive, assistive technology system that recognizes and classifies ADL. Using the Web, the system can provide information on current patient behavior that can help to quantify the patient's cognitive impairment, estimate the patient's self-dependency, and facilitate formal and informal care of Alzheimer's disease patients living alone. We present here the results of the new Web-based, non-intrusive wireless sensor system and an ADL classifier. We hope that the outcome of our study, together with further investigation and research, will justify further interventional studies.

Methods

Overview

The components of this novel assistive technology system are (1) a number of sensors that are distributed in every room and (2) a central computer unit (CCU). The collected data are stored and processed by an ADL classifier based on a rule-based forward chaining inference engine. Eight ADL, such as sleeping, grooming, toileting, getting ready for bed, cooking, eating, watching TV, and seated activity were allotted for the ADL classifier to determine. The ADL were chosen in consultation with medical experts. Table 1 shows the eight ADL in detail, along with a definition for each.

Table 1. The eight ADL in detail, with definitions.

ADL	Definition
Sleeping	Includes: Night rest, taking a nap (either in bed or on the couch). Excludes: Lying down (not sleeping) for recovery.
Grooming	Includes: Personal hygiene: Showering, toileting, shaving, brushing teeth, and styling as one activity. Excludes: Simple toileting and hand washing.
Toileting	Includes: Simple toileting with washing hands. Excludes: Other or additional personal hygiene.
Getting ready for bed	Includes: Personal hygiene before bedtime. Excludes: Pre-bedtime rituals.
Cooking	Includes: Preparing food in the kitchen. Excludes: Cutting pizza from delivery service, making popcorn, etc, making tea or coffee.
Eating	Includes: Having a meal (also delivered food). Excludes: Snacking (eg, while watching TV), having just a cup of coffee or a glass of water.
Watching TV	Includes: Watching TV with main focus on the TV. Excludes: Other activities while the TV is just on.
Seated activity	Includes: Sitting at a table or in an easy chair while reading, solving a puzzle, doing crosswords, embroidering, doing crafts, or listening to the radio. Excluding: Taking a nap.

Participants

Healthy participants were recruited by advertisements in local newspapers. They were assessed with a standardized paper-pencil test battery, which included the Montreal Cognitive Assessment (MoCA) [24,25], the Trail Making Test A and B (TMT-A, TMT-B) [26], and the Timed Up and Go test [27,28]. Exclusion criteria for the study were cognitive impairment (MoCA score < 26), or significant motor impairment (Timed Up and Go > 12 seconds). Participants sharing the house with others and not living alone were also excluded.

In total, 10 healthy mobile adults (6 women, 4 men; age 28-79 years) were included in the study. The study was carried out in accordance with the Declaration of Helsinki and was approved by the local ethics board. Written informed consent was obtained from all participants prior to inclusion. No compensation for participation was provided.

The Assistive Technology System

In each case, five sensors were mounted on a printed circuit board (PCB) capturing ambient values: temperature in °C (DS18B20, Dallas Inc), humidity in g/m^3 (SHT21P, SENSIRION), luminescence in lx (AMS302, Panasonic Inc), motion (binary) (EKMB1101111, Panasonic Inc), and acceleration in m/s^2 (ADXL345, Analog Device). Each PCB was assembled in polyvinyl chloride (PVC) housing (Figure 1) and was powered by a 2900 mA AAA primary cell.

In total, 50 sensors were assembled to 10 wireless sensor boxes (*l* x *w* x *h*=15mm x 30mm x 60mm, weight=80g), which digitalizes the analog environmental data and sends it to the CCU. A commercial available laptop, running customized Microsoft Windows 7, builds the CCU and served as data server with integrated Web link (Figure 1). The receiver unit (Figure 1), which is attached to the CCU, collected the data packages sent from all 10 sensor boxes. It comes with a digital received signal strength indication, an input sensitivity of 117 dBm, and a programmable transmitter output power up to 13 dBm. The CCU has the computing power to process the environmental data but also to store the data for further analysis.

The environmental data were acquired at a sampling rate of 0.2 Hz and assembled to data packages. Beside the five environmental values, each data package includes a handshake word composed of timestamp, date, node number, supply voltage, and a status word. The data packages were sent to the data server over an air link based on EZRadioPro-Technology (Silabs). Therefore, a low byte order marked bidirectional protocol placed on an 868 MHz carrier transmitted by frequency modulation was implemented. Even-parity error handling and frame collision detection was implemented.

Figure 1. Wireless sensor box with housing (left) and an inside view of the same sensor box (middle) displaying PCB board with environmental sensors and primary cell; receiver unit with antenna (right) is connected to the central computer unit serving as a data server and Web link (between the two devices a bidirectional data transmission is established).

System Set-up

The assistive technology system was installed in the home of 10 healthy subjects. Figure 2 shows the floor plan of a sample 2-bedroom apartment. In each apartment, each compartment (room) was fitted with one sensor box, placed at a height of approximately 2 meters, facing towards the middle of the room. Additional sensor boxes were placed in the kitchen (on the fridge door) and in the bathroom (on the flush handle).

Installation duration for the entire system depends on the apartment layout and varied from 15-30 minutes. Once set up and initialized, the system continuously recorded the five ambient environmental values autonomously. Activities of the 10 subjects were recorded for 20 days each. The subjects were told to ignore the system as much as possible and act naturally. They were also asked to touch the sensor boxes only if absolutely necessary. The system was uninstalled after 20 days of continuous measuring.

Figure 2. Example setup installed in a 2-bedroom apartment (sensor boxes indicated with red circles).

Subject Log Book

For evaluation purposes, a wireless protocol device, built in a housing with a wearable belt clip, was provided to the subjects (Figure 3). The protocol device was fitted with switches, each corresponding to a specific ADL. The ten switches provided on the protocol box were "Sleeping", "Grooming", "Toileting", "Getting ready for bed", "Cooking", "Eating", "Watching TV", "Seated activity", " receiving visitors", and "cleaning". Of the ten switches, one activity (cleaning) and one state (receiving visitor) were not included in the analysis. All subjects were instructed to record their activities throughout the day for the duration of the 20 days measurement. For this purpose, they were asked to flick the corresponding switch, labeled on the protocol device case, during the performance (exact duration) of the given ADL. The protocol device was intended to be worn during the day. Only if the subjects were disturbed by wearing the device (eg, during nighttime or when grooming) were they allowed to take it off.

In addition, participants were provided with a paper-pencil log book, to record every wrongly stated, forgotten, or missed ADL.

Figure 3. Wireless protocol device worn by participant (switches on top of the box are labeled with 8 different activities such as "watching TV" or "sleeping").

Data Processing

First, two different sorting methods were used sequentially to increase the sorting efficiency of the acquired data (Figure 4). A Bucketsort algorithm [29] was first applied to the data, which were labeled with a compartment identifier by the system during acquisition. Each data value was allotted to the compartment (room) where the value was captured. Metaphorically, each compartment was represented by one bucket. The data in each bucket were then sorted with a Radixsort [30] algorithm. The five bits considered by the Radixsort are related to the intra-day-timeline. This sorting of data is needed to line up the measured values chronologically.

In summary, the data are bijectively matched to their origin compartment by the Bucketsort, and then lined up in a chronological order by Radixsort. A mathematical description regarding the efficiency of the two sort algorithms is presented in Multimedia Appendix 1, and the open source code is provided in Multimedia Appendix 2.

In the second step, the data were classified using an ADL classifier (Figure 5). The classification of the ADL is based on the assumption that each subject follows a daily routine, where specific patterns with nearly the same duration and course occur throughout each day [31]. Considering the fact that numerous data values are accumulating during the measurements, the classifier was implemented as a rule-based inference engine. The concept is ideal to handle numerous data, as it is widely

XSL•FO
RenderX

used in the field of very-large-scale integration [32]. The theoretical concept was adapted to match the needs of the ADL classifier. It consists of (1) a database, (2) a rule-repository, and (3) a forward chaining inference engine. The database (1) holds all the data sorted upfront by the Bucketsort and Radixsort algorithms, but also all the classified ADL so far (historical data). The rule-repository (2) provides the forward chaining inference engine with a set of parameterized behavioral knowledge (the parameterization was done in cooperation with our medical experts.) A parser translates the parameterized behavioral knowledge into a look-up table disposable in the random access memory. The (3) forward chaining inference engine charges all available facts according to the given rules. The rules were defined manually and were the same for all subjects. The daily routine itself (ie, the time of the activity) was not considered within the rules; rather, the rules were applied to the daily routine resulting from the specific behavior pattern throughout the day. Nevertheless, the forward chaining inference engine needs some conflict resolution strategy to decide which information is the most important to process first and in which order the rest of the information has to be taken into account. Going the other way, the forward chaining inference engine checks which condition information must be fulfilled to state the given information as one specific ADL.

The used method is to obey the rules that the medical experts defined first. The forward chaining inference engine, on the other hand, works with three basic elements: (1) ambient value matrices, (2) rules, and (3) behavioral parameters. For each potential ADL, an ambient value matrix is composed of the sorted raw data. Each ambient value matrix is then compared to a set of rules defined upfront for each of the eight ADL. The rules are built under the premise of the behavioral parameter. If the ambient value matrix fits the set of rules, the activity is considered as the corresponding ADL.

For example, if one looks at "toileting", the matrix consists of the data from the five sensors in the bathroom and from the flush handle. Hence, the ambient value matrix shows a specific pattern associated with toileting. First, the forward chaining inference engine verifies the activity's duration for toileting. Further, the forward chaining inference engine verifies if the light conditions in the bathroom change, if the subjects sit on the toilet, if the subjects flushes, and if the humidity and the temperature do not change significantly. Having these prerequisite (rules) fulfilled, the activity is stated as "toileting". The mathematical formulas and the source code of the forward chaining inference engine are provided in Multimedia Appendix 3.

Depending on the complexity of the ADL to be classified, n-steps of iteration can be done. By processing the daily routine with the developed classifier algorithm, the system allots specific patterns to one of the 8 selected ADL. The output of the classifier (the conclusion) is the ADL a subject performed at a given time throughout the day. This information is stored in the database.

The calculation and algorithms were implemented in Matlab R2007b (MathWorks, Inc).

Figure 4. Combined sort algorithm consists of a Bucketsort and a Radixsort algorithm applied to the data.
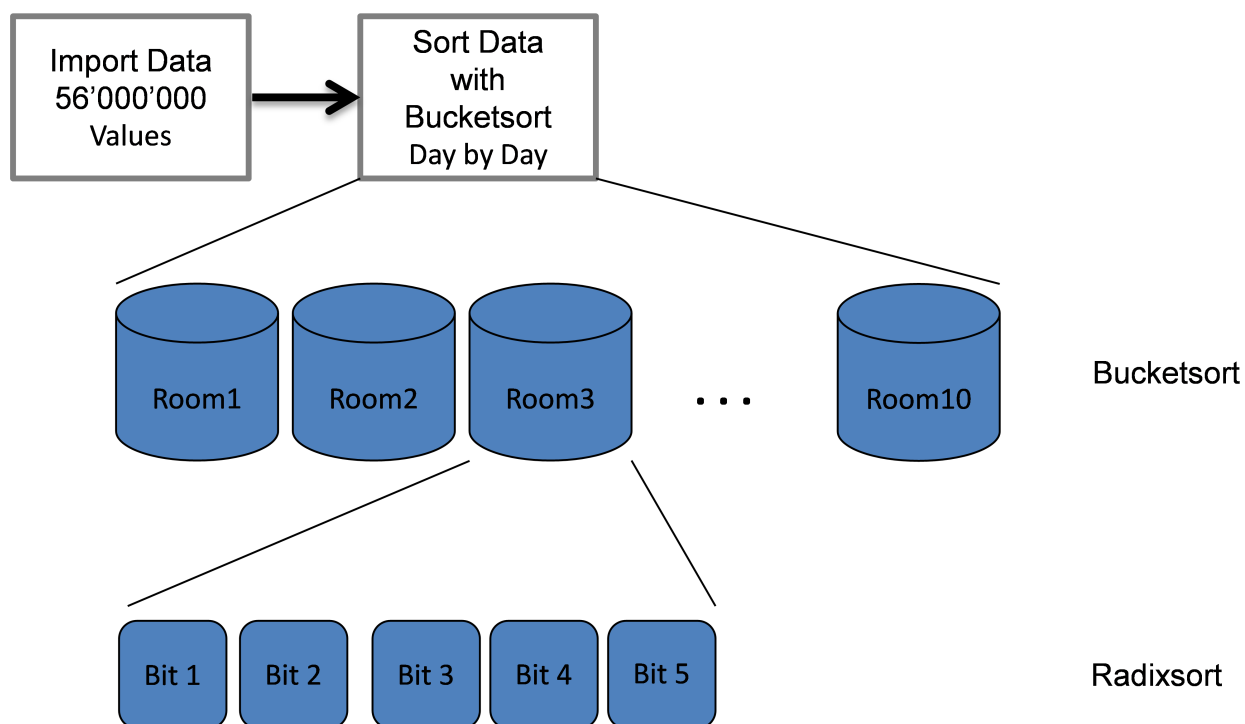
Figure 5. Block diagram of the ADL classifier algorithm based on a forward chaining inference engine.
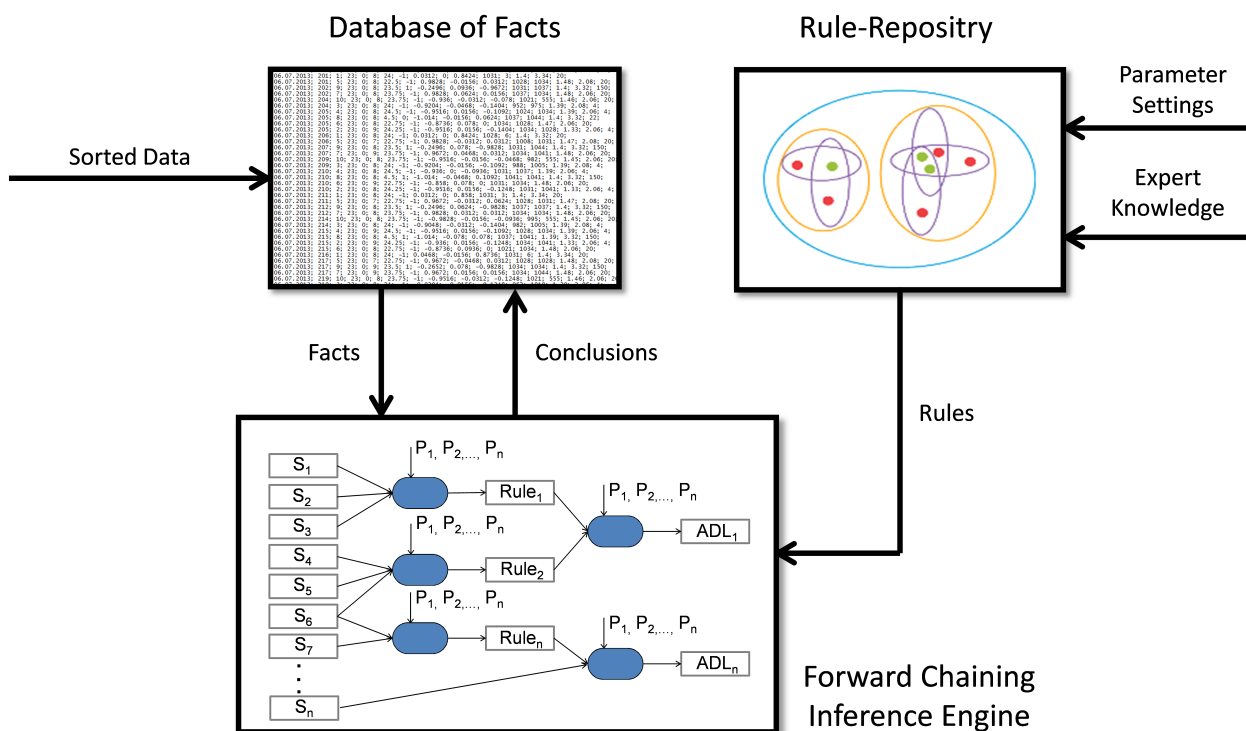
Data Analysis

The subject log book was updated by superimposing the data of the wireless protocol device with the paper-pencil log book.

The systems' performance (sensitivity and specificity) was then calculated by correlating the output of the ADL classifier with the ADL protocols (Figure 6).

Figure 6. Correlation of the ADL classifier output with the ADL protocols.
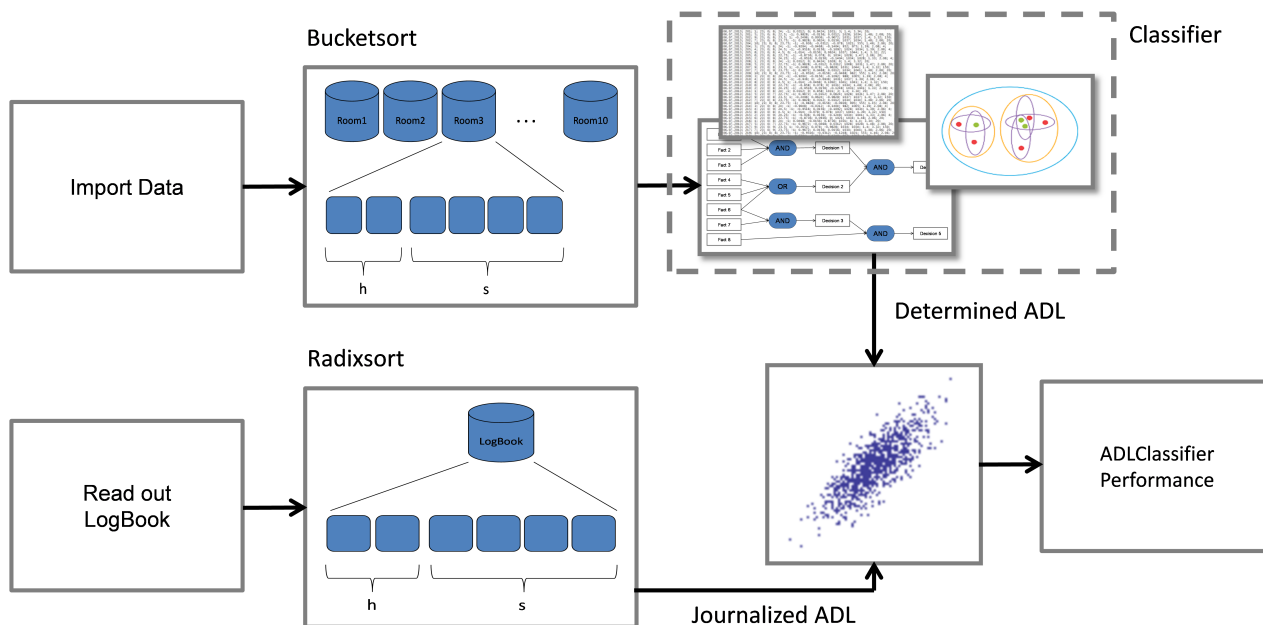
Data Collection With an Alzheimer Patient

In addition to the 10 healthy subjects, one dementia patient was measured for explorative purpose. This was done to understand the acceptance of the system in the target group of dementia patients and also to estimate the further clinical implications. The measurement was performed similarly in all respects to the measurement in the healthy subjects. The patient was an 84-year-old female. She was receiving home care twice a week for 1 hour, and she was on her own for the rest of the time.

Results

Demographics

Table 2 shows the demographics of 10 healthy participants (6 women, 4 men; mean age 48.8 years; SD 20.0 years; age range

28-79 years) included in the study. The test performance in MoCA (score range 27-30), TMT-A (time range 15.3-81.4 seconds), TMT-B (time range 21.1-122.6 seconds), and Timed Up and Go test (time range 6.9-10.3 seconds), of all the participants were in a normal, non-pathological range.

Table 2. Demographics of 10 healthy participants included in the study.

Characteristics		Value
Gender, n		
	Male	4
	Female	6
Age, years		
	Minimum	28.0
	Maximum	79.0
	Mean	48.8
	SD	20.0
MoCA Score (Maximum = 30)		
	Mean	29.1
	SD	1.1
Timed Up & Go		
	Mean	8.2
	SD	1.3
Trail Making Test (TMT)		
	Mean TMT A	39.1
	SD TMT A	20.0
	Mean TMT B	62.6
	SD TMT B	32.3
Measured time, days		
	Mean	20.0
	SD	0.0

Data Analysis and Activities of Daily Living Classifier Performance

In total, 33,939,441 data packets (543,031,056 environmental values) were captured correctly, and 160,269 data packets (0.47%) were lost due to transmission error. Thus, an overall reliability of 99.53% was achieved.

In total, 1317 ADL were performed by the participants, 1211 ADL were classified correctly, and 106 ADL were missed. This leads to an overall sensitivity of 91.27% and a specificity of 92.52% (Table 3). Each subject performed 134.8 ADL on average (SD 75).

Table 3. Results of the ADL classifier.

ADL	N	Classified correctly, n	Missed, n	Sensitivity, %	Specificity, %
Sleeping	173	161	12	93.64	85.77
Grooming	135	127	8	94.07	96.98
Toileting	307	291	16	94.79	91.54
Getting ready for bed	105	97	8	92.38	94.48
Cooking	70	59	11	84.29	90.92
Eating	90	78	12	87.78	94.83
Watching TV	322	300	22	93.17	90.63
Seated activity	171	151	20	90.06	94.98
Total	1317	1211	106	91.27	92.52

Behavior Pattern of an Alzheimer Patient Versus a Healthy Subject

Figure 7 shows the activity pattern for the 84-year-old female Alzheimer patient (MoCA score=23, Timed Up and Go=19.8 seconds, Trail Marking Test A=84.3 seconds, Trail Marking Test B=146 seconds).

The activity pattern of a healthy female (age=79, MoCA score=29, Timed Up and Go=12.7 seconds, Trail Marking Test A=62.2 seconds, Trail Marking Test B=95.1 seconds) is shown in Figure 8.

Figures 7 and 8 show the duration and points in time of the performed ADL of the Alzheimer patient versus the healthy subject.

Figure 7. Activity map of 84-year-old female Alzheimer patient.
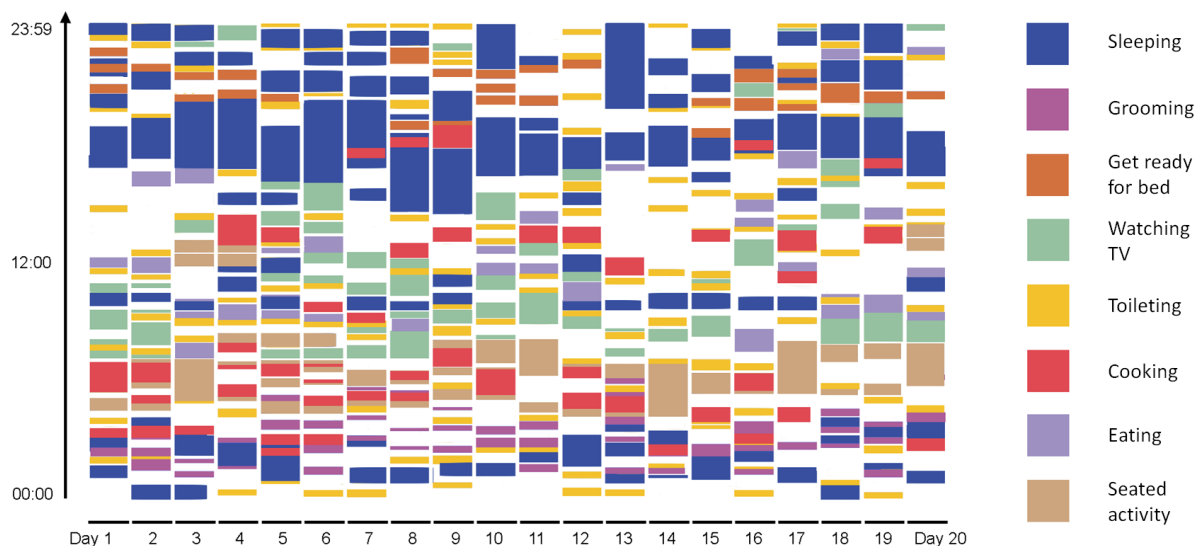
Figure 8. Activity map of healthy 79-year-old female.
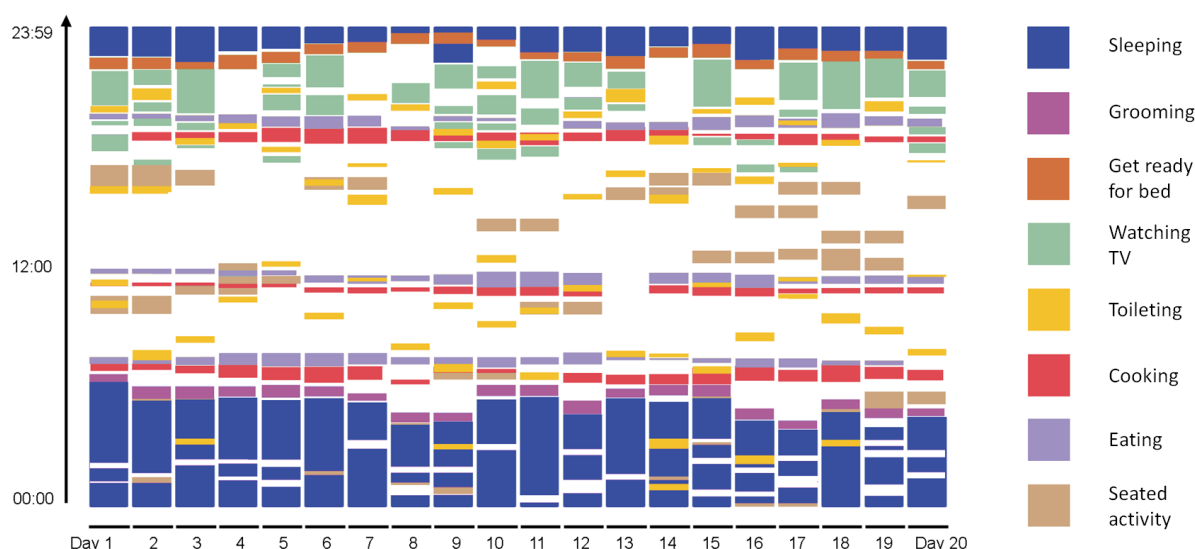
Discussion

Principal Findings

The feasibility and reliability of the newly developed non-intrusive sensor system (hardware and the operating software) could be proven in the field. The results show that daily living evokes variation in ambient values that can be captured by the sensor system, leading to ADL specific data patterns. Hence, it is possible to assign sensor data patterns to specific activities, whereby the most relevant sensor in the

detection of ADL is the movement sensor. By analyzing the retrieved data, it became possible to identify eight ADL.

Other researchers have conducted similar studies. Bang et al [12] used a set of pressure sensors, passive infrared sensors, and a worn accelerometer to determine ADL. Depending on the specific ADL, they achieved an initial accuracy of 93.27% to 96.47%. Fleury et al [16] used video cameras and wearable kinetic sensors, door contacts, and microphones to acquire data. The data were processed by a support vector machine to determine specific ADL. They achieved a sensitivity of 97.80%.

Both results are slightly better than what we found—at the cost of more intrusive sensor systems.

The ADL classification performance of the system relied on the protocol compliance of the subjects during the measurement. We asked the subjects to record their ADL very carefully with the electronic protocol device and provided them furthermore with paper-based log books to note any errors or wrong manipulations. Overall, it is possible that subjects forgot to record an activity or recorded a wrong activity, which negatively affects the classification performance numbers. Hence, the reported performance of the ADL classifier expressed by sensitivity and specificity numbers actually represents worst-case scenario.

A limitation of our system is that the sensors need a clear view of the entire room and must not be covered. This is can be solved by proper placement of the sensors in the patient's home.

As expected, the classifier showed better performance for some ADL (eg, grooming) than it did for others (eg, sleeping). This is due to the fact that some ADL are not exclusively room related and evoke only slightly different ambient values. Such data pattern differs only poorly among one another, which has also been found by other researchers [12,16,33].

The system uses wireless technology with small, discrete shaped sensor boxes. Hence, no hardware or tools are needed for the installation and the set-up can be done in less than 30 minutes. All parts were built under consequent low-power conditions, which keeps the system in service for 6 months straight. During this period, no maintenance is needed. Once the system is set up, the functioning of each sensor box can be monitored at any time. This leads to a high degree of transparency regarding the transmitted, lost, and compromised data (reliability of 99.53%). Furthermore, data can be bijectively matched to its origin to take statistical advantage of paired data, which ensures valid data throughout the measurement. Moreover, by foregoing the use of cameras and microphones, the privacy of the subjects can be guaranteed at any time of the experiment. Furthermore, due to its Web link, the system can provide discrete information about the patient's ADL performances in real-time, but also historical and statistical information about the progression of the disease.

When it comes to measurement in dementia patients, we proved feasibility in one female Alzheimer patient. The acceptance of the system was good, and no significant incident or failure occurred during the measured period. The patient was not disturbed by the system.

The activity map of the patient shows significant difference to the activity map of the healthy subject. While in the healthy subject, a daily structure is observable, much less structure can be found in the Alzheimer patient. Comprehensive analysis of further datasets of dementia patients can provide valuable clinical information, especially by comparing it to datasets of healthy patients.

These preliminary results obtained from one Alzheimer patient are promising in anticipation of further research and the clinical implication of the system.

Conclusions

The non-intrusive wireless sensor system can be used to acquire environmental data essential for the classification of ADL. By analyzing the retrieved data, it is possible to distinguish and assign data patterns to subjects' specific activities and to identify eight different ADL. Thanks to the Web-based technology, the system has a high potential to improve care and provides valuable information about the patient in real-time.

Based on the results of this study, there are plans to install the system in the home of Alzheimer patients, suffering from moderate to severe dementia. The data of these patients will be compared to the data of the healthy subjects to study and analyze the different behavior patterns.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Bucketsort and Radixsort formulas.

[PDF File (Adobe PDF File), 104KB - jmir_v16i7e175_app1.pdf]

Multimedia Appendix 2

Source code.

[ZIP File (Zip Archive), 19KB - jmir_v16i7e175_app2.zip]

XSL•FO

RenderX

Multimedia Appendix 3

Forward Chaining Inference Engine formulas.

[PDF File (Adobe PDF File), 104KB - jmir_v16i7e175_app3.pdf]

References

1.   Pollack ME. Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairment. AI magazine. Vol.26, No.2 2005:9. [doi: 10.1609/aimag.v26i2.1810]

2.   WHO. Dementia cases set to triple by 2050 but still largely ignored. WHO Media Centre 2012:1.

3.   Juchli L. Pflege. Stuttgart: Thieme; 1998.

4.   Kraft E, Marti M, Werner S, Sommer H. Cost of dementia in Switzerland. Swiss Med Wkly 2010;140:E1-E8 [FREE Full text] [doi: 10.4414/smw.2010.13093] [Medline: 22250014]

5.   Pincus T, Summey JA, Soraci SA, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum 1983 Nov;26(11):46-53. [Medline: 6639693]

6.   Wade DT, Collin C. The Barthel ADL Index: a standard measure of physical disability? Int Disabil Stud 1988;10(2):64-67. [Medline: 3042746]

7.   Arcelus A, Jones MH, Goubran R, Knoefel F. Integration of smart home technologies in a health monitoring system for the elderly. In: IEEE Conference. 2007 Presented at: IEEE Conference; 21-23 May 2007; Niagara Falls, Ont p. 820-825. [doi: 10.1109/AINAW.2007.209]

8.   Popescu M, Chronis G, Ohol R, Skubic M, Rantz M. An eldercare electronic health record system for predictive health assessment. In: IEEE Conference. 2011 Presented at: IEEE Conference; 13-15 June 2011; Columbia, MO. [doi: 10.1109/HEALTH.2011.6026742]

9.   Rowe M, Lane S, Phipps C. CareWatch: A Home Monitoring System for Use in Homes of Persons With Cognitive Impairment. Top Geriatr Rehabil 2007 Jan 1;23(1):3-8 [FREE Full text] [Medline: 20454554]

10.   Wood A, Virone G, Doan T, Cao Q, Selavo L, Wu Y, et al. ALARM-NET: Wireless sensor networks for assisted-living and residential monitoring. University of Virginia, Department of Computer Science 2006:1-14.

11.   Almudevar A, Leibovici A, Tentler A. Home monitoring using wearable radio frequency transmitters. Artif Intell Med 2008 Feb;42(2):109-120. [doi: 10.1016/j.artmed.2007.11.002] [Medline: 18215512]

12.   Bang S, Kim M, Song SK, Park SJ. Toward real time detection of the basic living activity in home using a wearable sensor and smart home sensors. In: Conf Proc IEEE Eng Med Biol Soc. 2008 Presented at: IEEE Conference; 20-25 Aug. 2008; Vancouver, BC p. 5200-5203. [doi: 10.1109/IEMBS.2008.4650386]

13.   Sixsmith A, Johnson N. A smart sensor to detect the falls of the elderly. IEEE Pervasive Comput 2004 Apr;3(2):42-47. [doi: 10.1109/MPRV.2004.1316817]

14.   Suzuki T, Murase S, Tanaka T, Okazawa T. New approach for the early detection of dementia by recording in-house activities. Telemed J E Health 2007 Feb;13(1):41-44. [doi: 10.1089/tmj.2006.0033] [Medline: 17309353]

15.   Tapia EM, Intille SS, Haskell W, Larson K, Wright J, King A, et al. Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. 2007 Presented at: 11th IEEE International Symposium on Wearable Computers; 11-13 Oct. 2007; Boston, MA p. 37-40. [doi: 10.1109/ISWC.2007.4373774]

16.   Fleury A, Vacher M, Noury N. SVM-based multimodal classification of activities of daily living in Health Smart Homes: sensors, algorithms, and first experimental results. IEEE Trans Inf Technol Biomed 2010 Mar;14(2):274-283. [doi: 10.1109/TITB.2009.2037317] [Medline: 20007037]

17.   Hopf J, Kistler R, Klapproth A. iHomeRescue - Ein multimedial vernetztes Sicherheitskonzept für ein sicheres Wohnen im Alter. 2011 Presented at: AAL - 4. Deutscher Kongress; 25-26 Jan. 2011; Berlin.

18.   Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980 Dec;20(6):649-655. [Medline: 7203086]

19.   Hanson E, Magnusson L, Arvidsson H, Claesson A, Keady J, Nolan M. Working together with persons with early stage dementia and their family members to design a user-friendly technology-based support service. Dementia 2007 Aug 01;6(3):411-434. [doi: 10.1177/1471301207081572]

20.   Beaudin JS, Intille SS, Morris ME. To Track or Not to Track: User Reactions to Concepts in Longitudinal Health Monitoring. J Med Internet Res 2006 Dec 7;8(4):e29. [doi: 10.2196/jmir.8.4.e29]

21.   Rialle V, Ollivet C, Guigui C, Hervé C. What do family caregivers of Alzheimer's disease patients desire in smart home technologies? Contrasted results of a wide survey. Methods Inf Med 2008;47(1):63-69. [Medline: 18213430]

22.   Demiris G, Rantz M, Aud M, Marek K, Tyrer H, Skubic M, et al. Older adults' attitudes towards and perceptions of "smart home" technologies: a pilot study. Med Inform Internet Med 2004 Jun;29(2):87-94. [doi: 10.1080/14639230410001684387] [Medline: 15370989]

23.   Kutz D, Shankar K, Connelly K. Making Sense of Mobile-and Web-Based Wellness Information Technology: Cross-Generational Study. Journal of Medical Internet Research 2013;15(5):e83. [doi: 10.2196/jmir.2124]

XSL•FO

RenderX

24. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]

25. Freitas S, Simões MR, Marôco J, Alves L, Santana I. Construct Validity of the Montreal Cognitive Assessment (MoCA). J Int Neuropsychol Soc 2012 Mar;18(2):242-250. [doi: 10.1017/S1355617711001573] [Medline: 22115097]

26. Reitan RM, Wolfson D. The Trail Making Test as an initial screening procedure for neuropsychological impairment in older children. Arch Clin Neuropsychol 2004 Mar;19(2):281-288. [doi: 10.1016/S0887-6177(03)00042-8] [Medline: 15010091]

27. Bischoff HA, Stähelin HB, Monsch AU, Iversen MD, Weyh A, von Dechend M, et al. Identifying a cut-off point for normal mobility: a comparison of the timed 'up and go' test in community-dwelling and institutionalised elderly women. Age Ageing 2003 May;32(3):315-320 [FREE Full text] [Medline: 12720619]

28. Podsiadlo D, Richardson S. The timed Up & Go: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [Medline: 1991946]

29. Corwin E, Logar A. Sorting in linear time - variations on the bucket sort. Journal of Computing Sciences in Colleges 2004;20:197-202.

30. Blelloch GE, Leiserson CE, Maggs BM, Plaxton CG, Smith SJ, Zagha M. A comparison of sorting algorithms for the connection machine CM-2. 1991 Presented at: ACM symposium on Parallel algorithms and architectures; 1991; New York, NY p. 3-16. [doi: 10.1145/113379.113380]

31. Virone G, Alwan M, Dalal S, Kell SW, Turner B, Stankovic JA, et al. Behavioral patterns of older-adults in assisted living. Information Technology in Biomedicine 2008 May;12(3):387-398. [Medline: 18693506]

32. Grifin NL, Lewis FD. A rule-based inference engine which is optimal and VLSI implementable. Architectures, Languages and Algorithms 1989:246-251. [doi: 10.1109/TAI.1989.65327]

33. Bouten CV, Koekkoek KT, Verduin M, Kodde R, Janssen JD. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. Biomedical Engineering 1997 Mar;44(3):136-147. [doi: 10.1109/10.554760] [Medline: 9216127]

Abbreviations

ADL: activities of daily living
CCU: central computer unit
MoCA: Montreal Cognitive Assessment
PCB: printed circuit board
PVC: polyvinyl chloride
TMT A & B: Trail Making Test A and B

Original Paper

Experiential Virtual Scenarios With Real-Time Monitoring (Interreality) for the Management of Psychological Stress: A Block Randomized Controlled Trial

Andrea Gaggioli[1,2], MS(Psycho), PhD(Psych); Federica Pallavicini[1], MS(Psycho), PhD(Psych); Luca Morganti[1], MS(Psycho); Silvia Serino[1], MS(Psycho); Chiara Scaratti[1], MS(Psycho); Marilena Briguglio[3], MD, PhD; Giulia Crifaci[3], BBiomedEng; Noemi Vetrano[3], MS(Psycho); Annunziata Giulintano[3], MS(Psycho); Giuseppe Bernava[3], M Comp Sc, PhD; Gennaro Tartarisco[3], MS Eng, PhD; Giovanni Pioggia[3], MS Eng, PhD; Simona Raspelli[1], MS(Psycho), PhD; Pietro Cipresso[1], MSc Econ, PhD; Cinzia Vigna[1], Media Spec; Alessandra Grassi[1], MS(Psycho), PhD; Margherita Baruffi[1], MS(Psycho), PsyD; Brenda Wiederhold[4], MBA, BCIA, PhD; Giuseppe Riva[1,2], MS(Psycho), PhD

[1]Applied Technology for Neuro-Psychology Lab, Istituto Auxologico Italiano, Milan, Italy
[2]Department of Psychology, Catholic University of Sacred Heart, Milan, Italy
[3]Institute of Clinical Physiology (IFC), National Research Council (CNR), Messina, Italy
[4]Virtual Reality Medical Institute, Brussels, Belgium

Corresponding Author:
Andrea Gaggioli, MS(Psycho), PhD(Psych)
Applied Technology for Neuro-Psychology Lab
Istituto Auxologico Italiano
Via Alessandro Magnasco 2
Milan, 20149
Italy
Phone: 39 0261911 ext 2892
Fax: 39 02619112892
Email: andrea.gaggioli@auxologico.it

Abstract

Background: The recent convergence between technology and medicine is offering innovative methods and tools for behavioral health care. Among these, an emerging approach is the use of virtual reality (VR) within exposure-based protocols for anxiety disorders, and in particular posttraumatic stress disorder. However, no systematically tested VR protocols are available for the management of psychological stress.

Objective: Our goal was to evaluate the efficacy of a new technological paradigm, Interreality, for the management and prevention of psychological stress. The main feature of Interreality is a twofold link between the virtual and the real world achieved through experiential virtual scenarios (fully controlled by the therapist, used to learn coping skills and improve self-efficacy) with real-time monitoring and support (identifying critical situations and assessing clinical change) using advanced technologies (virtual worlds, wearable biosensors, and smartphones).

Methods: The study was designed as a block randomized controlled trial involving 121 participants recruited from two different worker populations—teachers and nurses—that are highly exposed to psychological stress. Participants were a sample of teachers recruited in Milan (Block 1: n=61) and a sample of nurses recruited in Messina, Italy (Block 2: n=60). Participants within each block were randomly assigned to the (1) Experimental Group (EG): n=40; B1=20, B2=20, which received a 5-week treatment based on the Interreality paradigm; (2) Control Group (CG): n=42; B1=22, B2=20, which received a 5-week traditional stress management training based on cognitive behavioral therapy (CBT); and (3) the Wait-List group (WL): n=39, B1=19, B2=20, which was reassessed and compared with the two other groups 5 weeks after the initial evaluation.

Results: Although both treatments were able to significantly reduce perceived stress better than WL, only EG participants reported a significant reduction (EG=12% vs CG=0.5%) in chronic "trait" anxiety. A similar pattern was found for coping skills: both treatments were able to significantly increase most coping skills, but only EG participants reported a significant increase (EG=14% vs CG=0.3%) in the Emotional Support skill.

XSL•FO
RenderX

Conclusions: Our findings provide initial evidence that the Interreality protocol yields better outcomes than the traditionally accepted gold standard for psychological stress treatment: CBT. Consequently, these findings constitute a sound foundation and rationale for the importance of continuing future research in technology-enhanced protocols for psychological stress management.

Trial Registration: ClinicalTrials.gov: NCT01683617; http://clinicaltrials.gov/show/NCT01683617 (Archived by WebCite at http://www.webcitation.org/6QnziHv3h).

Introduction

The emerging convergence of technology and health care [1] is offering new methods and tools for mental health treatments [2-6]. An emerging trend is the use of virtual reality (VR) within the exposure-based protocols for anxiety disorders and posttraumatic stress disorders (PTSD) [7-11]. PTSD is more difficult to treat than other anxiety disorders. On one hand, in vivo exposure-based therapy is usually not possible. On the other, imaginal exposure requires that the patient recounts their traumatic experience in the present tense to the therapist—a behavior that patients try to avoid [12]. VR therapy allows exposure treatment even with patients who fail to improve with traditional imaginal exposure therapy [13-16]. Since the initial work by Barbara Rothbaum and her team [17,18], additional case studies [19-23] and clinical trials [22,24,25], including a randomized controlled clinical trial [26], have shown the efficacy of VR therapy in the treatment of PTSD.

One factor that may increase the likelihood of developing the symptomatology of PTSD is work-related stress. Previous studies have included stressful life events in the list of risk factors for PTSD, suggesting that experiencing chronic psychological stress may increase vulnerability to this anxiety disorder [27,28]. Indeed, chronic psychological stress may induce plasticity within the amygdala, which in turn may increase the risk of developing chronic anxiety states [29]. This abnormal change in the limbic neural circuitry may provoke a pathological anxiety response, leading to syndromes such as PTSD. Other studies have focused on the contribution of work-related stressors to PTSD [30,31]. For example, Laposa et al [30] found that interpersonal conflicts, inadequate support from superiors, and changing jobs were significantly associated with PTSD symptoms. Different training has been developed for managing psychological stress at work, both individually and organizationally focused [32,33]. Specifically, a recent review showed that individual interventions, like cognitive behavioral therapy (CBT), can improve individuals' mental health, while physical activity as an organizational intervention is more effective in reducing absenteeism [32]. On the basis of this evidence, we decided to focus our intervention on the individual.

However, until now, no systematically tested VR protocols have been available for the management of psychological stress. A preliminary attempt, developed by the US Army, was tested for the training of military medical professionals who are expected to take care of the wounded in very austere situations. This protocol included a technology-assisted relaxation training merging VR exposure sessions with relaxing videos with embedded English narratives guiding progressive muscle relaxation and controlled breathing [34].

According to Cohen et al [35], psychological stress occurs when people perceive that environmental demands tax or exceed their adaptive capacity. In this view, stressful experiences are conceptualized as person-environment transactions, whose results are dependent on the impact of the external stimulus. This is mediated by:

- The person's appraisal of the stimulus: when faced with a stimulus, a person evaluates the potential threat (primary appraisal). Primary appraisal is a person's judgment about the significance of a stimulus as stressful, positive, controllable, challenging, or irrelevant.
- The personal, social, and cultural resources available: facing a significant stimulus, a second appraisal follows, which is an assessment of the individual's coping resources and options. Secondary appraisals address what one can do about the situation.
- The efficacy of the coping efforts: if required by the appraisal process, the individual starts a problem management phase aimed at regulation of the external stimulus.

Stress Management Therapy can help counter effects of psychological stress. Usually various techniques are used including relaxation, interaction, biofeedback, and CBT methods. According to the Cochrane Database of Systematic Reviews [36-38], the best validated approach covering both stress management and stress treatment is the CBT approach. Typically, a CBT protocol (10-15 sessions) includes both problem-focused (eg, resource optimization and better planning) and emotion-focused (eg, relaxation training, use of emotional support) coping strategies. Initially, it includes in-session didactic materials and experiential exercises and out-of-session assignments (practicing relaxation exercises and monitoring stress responses).

The clinical intervention primarily focuses on (1) learning how to cope better with daily stressors (psychological stress) or traumatic events (PTSD), and (2) optimizing one's use of personal skills and social resources.

The trouble with managing psychological stress is that it is very personal. So the focus for assessment, prediction, and treatment has to be the situated experience of the individual. This result

is difficult to achieve using the available VR protocols for PTSD. From a clinical viewpoint, in these protocols VR provides a "closed" experience, separated from the emotions and behaviors experienced by the patient in the real world. In other words, VR exposure tries to change cognitive content per se, rather than changing the context in which cognitions are experienced [39]. The behavior of the patient in VR has no direct effects on the real-life experience. The emotions and problems experienced by the patient in the real world are not directly addressed in the VR exposure. Moreover, it focuses on patients' thoughts and behaviors but does not address social support and coping skills.

To overcome these issues, Riva et al [40-42] suggested the use of the "Interreality" paradigm (IR) that integrates assessment and treatment within a hybrid environment, bridging the physical and virtual world. The basic idea of the IR intervention is to bridge virtual experiences (fully controlled by the therapist, used to learn healthy behaviors and coping skills) with real experiences (the therapist can identify critical situations and assess clinical change). In the IR strategy, behavior in the physical world influences the experience in the virtual one, and behavior in the virtual world influences the experience in the real one. The current CBT approach can be described as imagining evokes emotions and the meaning of the associated feelings can be changed through reflection and relaxation. In IR, we introduce an alternative strategy, in which controlled experiences evoke emotions that result in meaningful new feelings, which can be reflected on and eventually changed through reflection and relaxation. This is achieved by using technology (virtual worlds, advanced sensors, and smartphones) to create a closed-loop approach for assessment and support.

The assessment is conducted continuously in the virtual and real worlds through the tracking of individuals' behavioral and emotional status over time, in the context of realistic task challenges. At the same time, the information is constantly used to improve both the appraisal and the coping skills of the patient through a conditioned association between effective performance state and task execution behaviors.

These features are integrated into two subsystems: the clinical VR platform (VR inpatient treatment, fully controlled by the therapist) and the mobile platform (mobile-based real world

support, available to the patient and connected to the therapist). Combined, these systems are able to provide (1) objective and quantitative assessment of symptoms using biosensors and behavioral analysis: monitoring of the patient behavior and of general and psychological status, early detection of symptoms of critical evolutions, and timely activation of feedback in a closed loop approach, and (2) decision support for treatment planning through data fusion and detection algorithms: the decision support system allows monitoring stress levels of the patient both during VR exposure (in the therapist's office) and in real-life situations (using the mobile phone), by generating reports that the therapist can access via a Web-based interface. The key features of IR are summarized in Figures 1 and 2.

Simply put, patients are continuously assessed in the virtual and real worlds by tracking their behavioral and emotional status in the context of challenging tasks (customization of the therapy according to the characteristics of the patient), and feedback is continuously provided to improve patients' skills (improvement of self-efficacy).

The potential clinical advantages of the IR strategy are (1) an integrated and quantitative assessment of the user's stress level using biosensors: the level of stress is continuously assessed in the virtual and in the real world by recording the participant's behavioral and emotional status for the decision support system, and (2) provision of warnings and motivating feedback to improve self-awareness, compliance, and long-term outcomes: on the basis of the decision support system, participants constantly receive feedback to improve their appraisal and coping skills in an entertaining and motivating manner both in clinical and mobile settings [40-42].

Starting from the above premise, the main goals of this study are (1) to define and develop an Interreality protocol for the management of psychological stress, and (2) to compare, within a controlled study, its efficacy with a similar non-technological protocol based on the CBT approach.

We hypothesize that the Interreality protocol is more effective than both standard CBT and a wait-list condition in (1) reducing the level of chronic "trait" stress, (2) reducing the perceived stress and improving quality of life, and (3) improving the coping skills of the individual.

Figure 1. Advantages of Interreality.

Figure 2. The Interreality paradigm for the management of psychological stress.
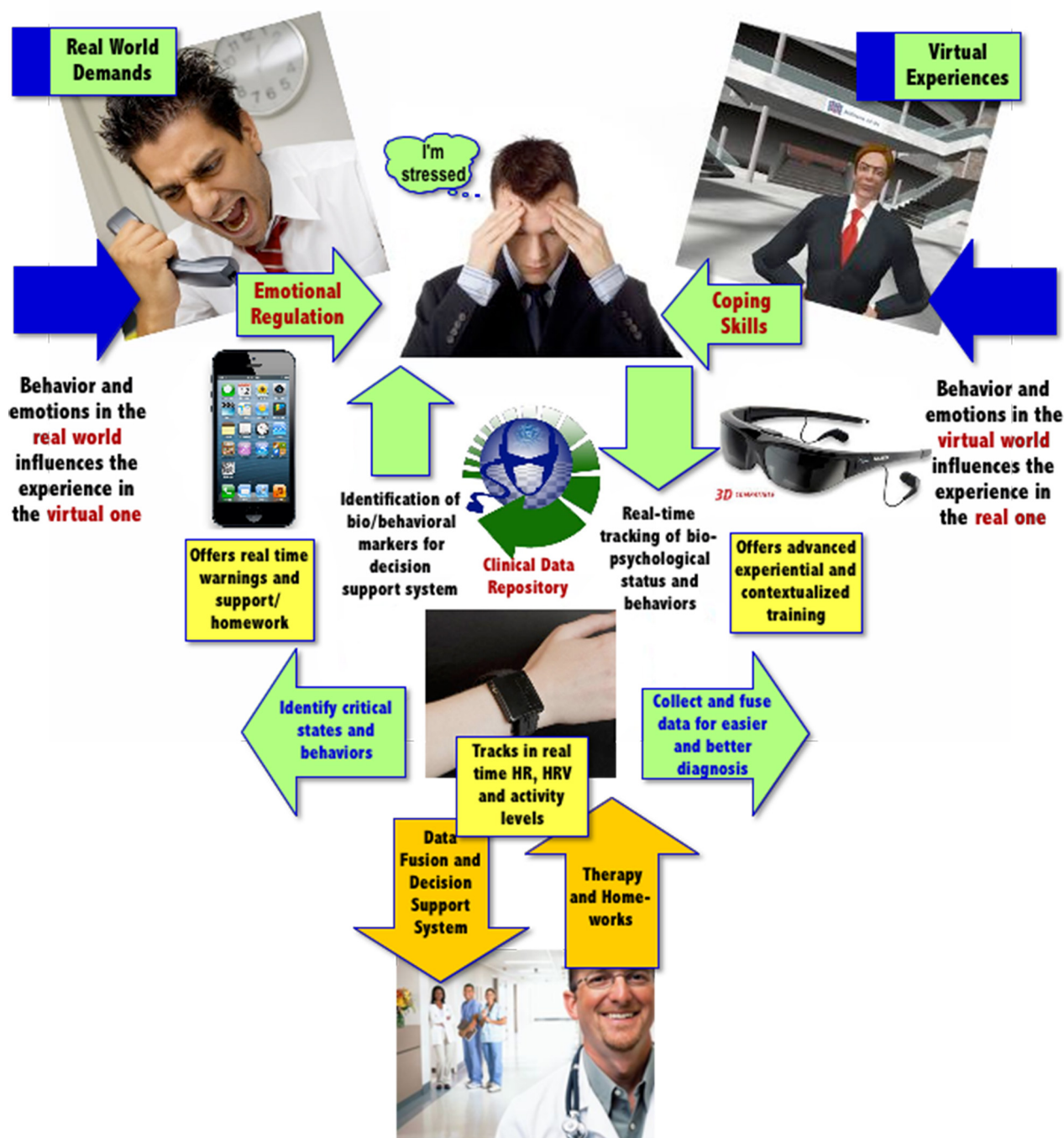
Methods

Recruitment

The study is designed as a multicentric randomized block controlled trial involving participants recruited from two different worker populations (teachers and nurses) that are highly exposed to psychological stress (the Consort flowchart is reported in Figure 3 and the electronic CONSORT-EHEALTH questionnaire [43] in Multimedia Appendix 1). Compared to a completely randomized design, this design reduces variability within treatment conditions and potential confounding (the variability within blocks is less than the variability between blocks), producing a better estimate of treatment effects [44].

A sample of high school teachers recruited in Milan (Block 1: n=95) and a sample of pediatric nurses recruited in Messina, Italy (Block 2: n=88) were seen for screening interviews for admission to the study. Samples were recruited between March 2012 and September 2013. Criteria for participation included the following: (1) a high level of perceived stress (≥7) as measured on a 10-item visual-analogue scale, (2) a high level of relevance of stress for personal health (≥7) as measured on a 10-item visual-analogue scale, (3) a low level of self-efficacy related to stress management (≤5) as measured on a 10-item

visual-analogue scale, (4) no DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, 4th edition, text revision) Axis I disorders as assessed by Mini-International Neuropsychiatric Interview (MINI) [45,46] during the clinical assessment, (5) aged 25-60 years, (6) no psychotherapy received for their psychological stress as assessed with a clinical interview, (7) no current psychiatric medications as assessed with a clinical interview, (8) no history of neurological diseases, psychosis, alcohol or drug dependence as assessed with a clinical interview, and (9) no migraine, headache, or vestibular abnormalities as assessed with a clinical interview.

Both males and females were included. In order to select participants, we decided to use only subjective indexes of stress and coping for this study, without measuring cortisol. The measures of the concentration of cortisol in blood, saliva, and urine are established methods for momentary assessments of the activity in the hypothalamic-pituitary-adrenocortical axis (HPA). If the cortisol levels become too high or too low for a longer period, a state of hyper- or hypocortisolism is present, and both are associated with stress-related disease. However, two recent studies by Barth [47] and Faresjo [48] suggest that cortisol levels are not very reliable indicators of stress. The first study [47] underlined the relevance of subjective evaluations in producing the negative effects of stress. The author found a reduced risk of coronary heart disease in stressed individuals

who neglected the subjective relevance of stress on health. The second study [48] demonstrated the inefficacy of measuring cortisol levels for assessing stress in subjects living in a stressful environment. The study found that living in a stressful economic and social environment produced a down-regulation of the HPA-axis with a suppression of cortisol levels.

In our study, 62 participants either did not fulfill inclusion criteria or refused to participate for other reasons (eg, time constraints). This is an unusual observation. However, it should be noted that participants experienced a high level of psychological stress. Chronic psychological stress alters the psychophysiological processes involved in cognitive appraisals and coping responses. Several types of coping strategy are commonly used to face the different demands associated with stressful events, and their effective use reflects variation in underlying cognitive appraisal. This refusal may therefore be read as their inability to appraise and cope with the stressful events.

All patients meeting the inclusion criteria in each block (B1=61, B2=60) were randomly assigned to the three groups: (1) the Experimental Group (EG): n=40, B1=20, B2=20, (2) the Control Group (CG): n=42, B1=22, B2=20, and (3) the Wait-List group (WL): n=39, B1=19, B2=20. All the participants signed an informed consent form before entering the study. The sample characteristics are shown in Table 1.

Table 1. Demographic parameters and baseline characteristics of the sample (mean and standard deviation).

Variables		EG	CG	W-L
Age		46.3 (7.7)	42.9 (10.5)	39.6 (9.7)
Years of education		17.9 (1.4)	17.3 (1.4)	18.3 (1.3)
PSM[a]		91.4 (25.8)	86 (19.5)	92.6 (25.8)
PSS[b]		21.1 (7.9)	18 (5.9)	19.4 (6.6)
STAI-Y2[c]		43.6 (11.2)	42.1 (11.1)	40.2 (9.8)
SWLS[d]		23.4 (6.2)	24.2 (7.2)	26.4 (7.0)
COPE-NIV [e]				
	Use of emotional support	27.8 (8.2)	29.9 (7.0)	29.4 (7.7)
	Positive attitude	27.9 (5.5)	29.1 (5.1)	31.8 (4.5)
	Problem focused	28.6 (6.0)	29.2 (7.4)	31.7 (4.6)
	Religious coping	19.8 (5.0)	21.7 (5.4)	23.8 (5.5)
	Denial	24.1 (5.9)	21.8 (4.7)	22.8 (3.8)

[a]Psychological Stress Measure.

[b]Perceived Stress Scale.

[c]State-Trait Anxiety Inventory Form Y-2.

[d]Satisfaction With Life Scale.

[e]Coping Orientation to the Problems Experienced—New Italian Version.

Baseline comparisons among the three groups showed only a difference in the age of the groups. The age was significantly higher in the EG and to a lesser extent in the CG than in the WL group. Of the EG participants, 57.5% were women (23/40) and 42.5% (17/40) were men, while the CG and WL included,

respectively: 71.4% women (30/42) and 28.6% men (12/42), and 51.3% (20/39) women and 48.7% men (19/39).

We also assessed participants' computer literacy at the start of the trial through a self-assessment scale with three values: "low", "medium", and "high". Results showed overall a medium level

of perceived individuals' technological abilities, and no

significant differences were found between groups.

Figure 3. Consort flowchart.
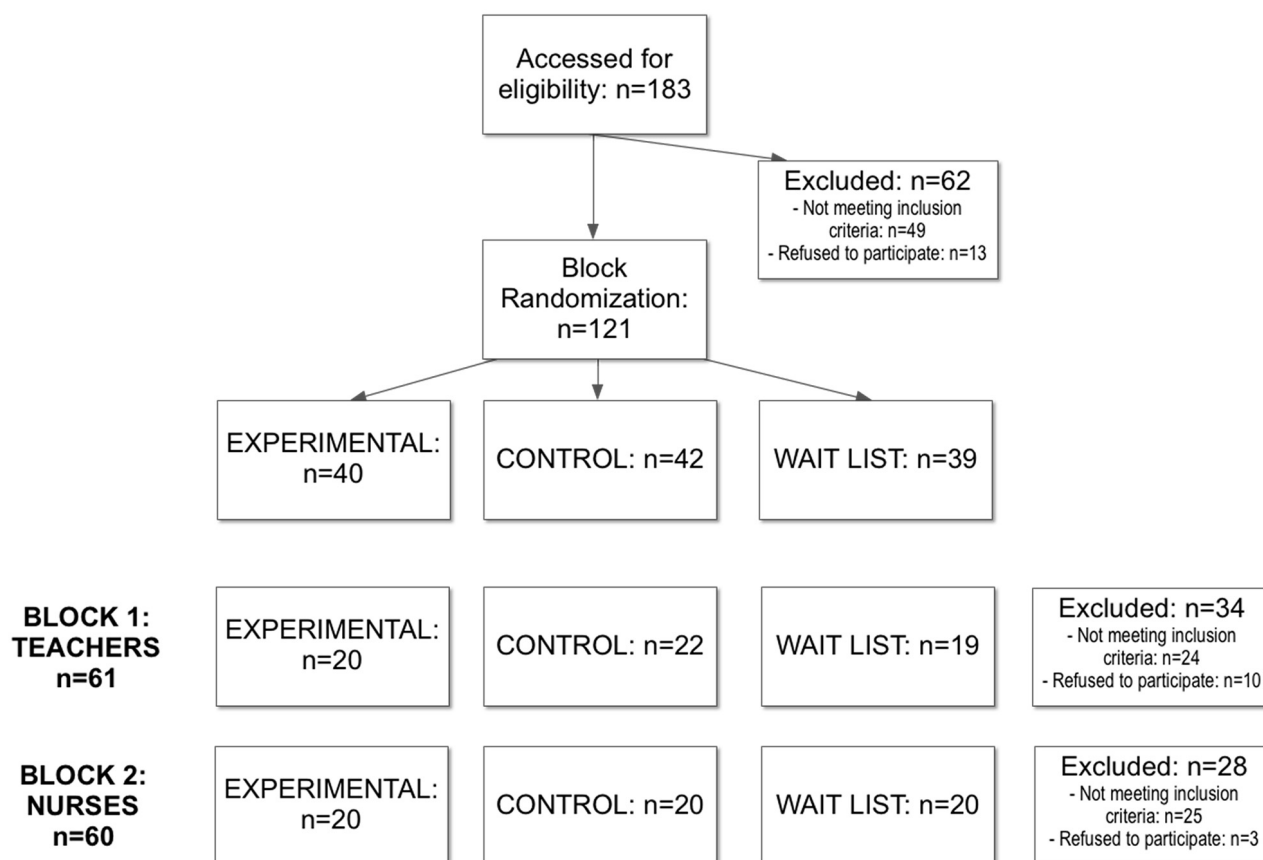
Ethics

The study was approved by ethical review board of Istituto Auxologico Italiano in Milan, Italy, and by the ethical review board of Azienda Ospedaliera Universitaria Policlinico "G. Martino" in Messina. The study was conducted according to the 1964 Declaration of Helsinki. All the participants signed an informed consent form that explained the goal of the treatment, its duration, the involvement of the patients, and for EG individuals, the possible side effects related to the extended use of immersive virtual environments (ie, ocular problems, disorientation and balance disturbances, and nausea).

Treatment Protocols

The treatment protocols were based on 5 weeks of stress management training (2 sessions per week) following the "stress management program" by Kaluza [49] and on the "stress inoculation training" by Meichenbaum [50].

Stress Management Training (SMT) is a short, focused, and individualized intervention to improve individual coping with stress at workplace [49]. A meta-analysis of 36 studies showed the efficacy of SMT in reducing negative mood states (ie, anxiety, depression) [51]. The Stress Inoculation Training (SIT) [50,52] is a validated short, semistructured, and active approach to manage psychological stress (for a review, see [53]). The effectiveness of SIT has been evaluated in different contexts. In the clinical setting, SIT has proven an effective means of

helping patients face particularly strenuous conditions [54-56]. In the occupational environment, SIT has been successfully applied to support employees in managing stressful situations [57,58] and to help athletes manage anxiety and improve performance [11,59].

The treatment block for teachers (Block 1: n=61) was offered by therapists from Istituto Auxologico Italiano both in the Istituto Scientifico Ospedale San Luca, Milano, Italy and in the schools where teachers worked. The treatment block for nurses (Block 2: n=60) was offered by therapists from the Azienda Ospedaliera Universitaria Policlinico "G. Martino" at the Pervasive Healthcare Center, a clinical research center of the Institute of Clinical Physiology.

The IR and CBT treatments were administered by clinical psychologists with extensive experience in stress management techniques. Detailed manuals were prepared to facilitate adherence with the treatment protocols. Clinical reports were checked for compliance at monthly supervision meetings. A high level of protocol adherence was reported by therapists. The difference between the treatment protocols offered by each block is detailed below (see also [41]).

Experimental Group

Overview

The experimental group used an Interreality protocol based on the following technologies (see Figure 4 for details).

Figure 4. Technologies used by the experimental group.
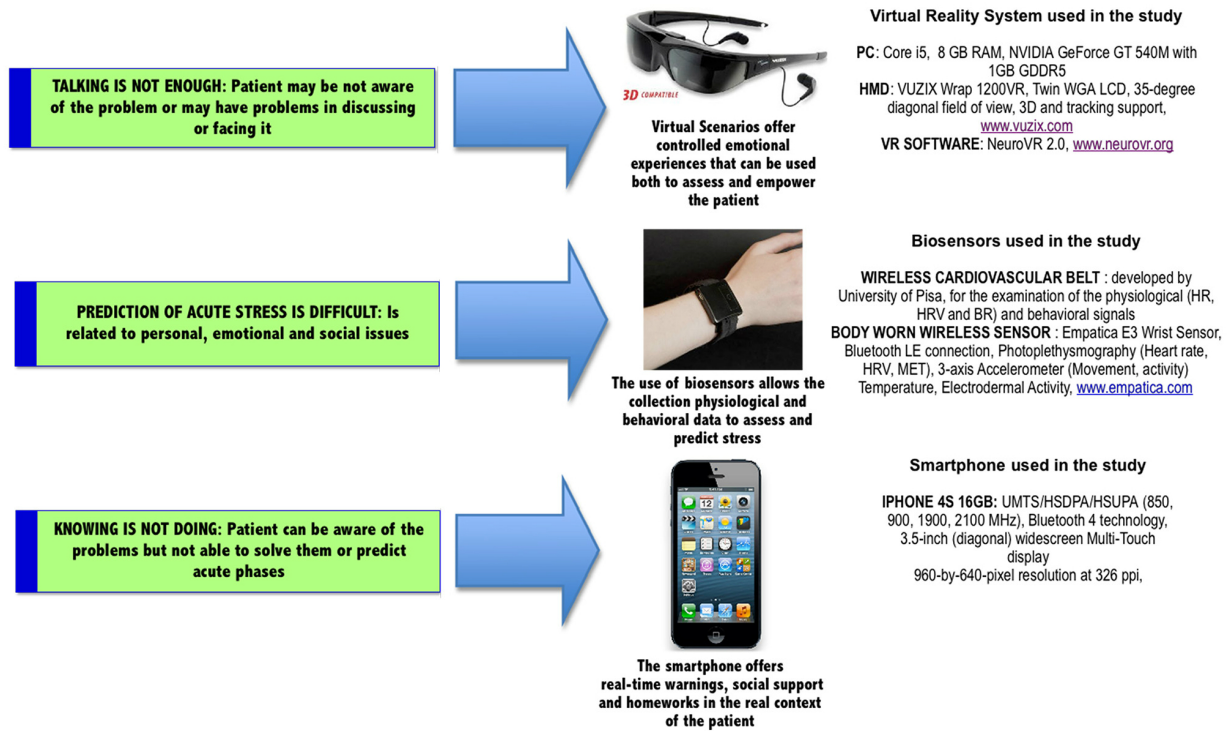
3D Virtual Scenarios (in the Therapist's Office)

Immersive Role-Playing Scenarios Where the Individual Interacts With Potentially Stressful Experiences

According to a literature review and the results of a qualitative analysis, different virtual stressful scenarios were generated for both teachers and nurses (see Table 2). The stressful scenarios were played by real actors, video-recorded, and included in the virtual environments (using the free virtual reality platform NeuroVR 2 [60-62]) after post-production (for a demonstration of the procedure, see the video in Multimedia Appendix 2 and Figure 5).

Figure 5. A Virtual Reality treatment session of the Interreality trial.

Immersive Natural Scenarios Used to Learn Specific Relaxation Techniques

In recent years, VR has been used in different clinical protocols to facilitate relaxation processes in stressed or anxious subjects [63-65] by visually presenting relaxing scenes [34,66]. The IR relaxation environments were created on the basis of similar virtual relaxing environments validated in previous studies [34,64,67-69]. They included a beach, a lake with a waterfall, a campfire in a mountain resort, and a desert oasis.

Table 2. The different virtual stressful scenarios.

Virtual scenarios for teachers	Virtual scenarios for nurses
Workload	Managing the patients' relatives
Class management	Managing patients' complaints
Coping with parent's criticism	Managing a medical emergency situation
Relationship with boss	Relationship with colleagues
Coping with parents' handling efforts	Managing relatives'/caregivers anxiety
Relationship with co-workers	Distribution of work tasks
Conflict management	Patient-doctor communication
	Managing patient's anxiety
	Unsuccessful collaboration/communication with colleagues
	Discussions among medical doctors

Biosensors (Personal Biomonitoring System)

Overview

All elements of technology, including smartphone and biosensors, were loaned to the EG participants free of charge. Moreover, all the required data connection fees for real-time stress monitoring were paid for by the trial. All the participants received 1-week group technology training plus personalized support sessions if needed.

Biofeedback (in the Therapist's Office)

Typically, 3D virtual worlds are a closed experience and do not reflect in any way the real activity and status of the users. In IR, physiological sensors (heart rate and heart rate variability) are used to track the emotional/health status of the individual and to influence their experience in the virtual world. To improve the efficacy of the relaxing environment, some features of the experience (eg, the size of the fire or the waterfall flow rate) were driven by the emotional status of the patient as measured by biosensors (heart rate or heart rate variability).

Physiological Data Recording (Outside the Therapist's Office)

To assess the level of contextual stress, each individual was provided a body-worn wireless sensor (EMPATICA E3 wrist sensor) that was able to record and transmit psychophysiological (heart rate and heart rate variation) and activity data in real time.

Mobile Phone (Outside the Therapist's Office)

Stress Tracking

The data received from the wireless sensors were assessed in real time by a decision support system (the description of the system that was used can be found in [70,71]). This system provides the user with a graphical representation of the current stress level experienced and allows them to check the history of stress levels variations experienced at different timescales (eg, day, week, month).

Contextualized Homework

According to the performance achieved in the therapist's office and level of stress assessed by the decision support system, the individual was able to experience on the smartphone different guided relaxation and biofeedback virtual experiences (non-immersive) similar to the ones experienced in the therapist's office.

Schema

Assessment Session (Session 1)

The session started with a discussion with the clinician about the assessment week. Then, after a brief introduction to the specific content of this session, the psychometric questionnaires were administered for the first time (see below), and the physiological baseline of the participant was recorded for 3 minutes. To measure the psychological variations occurring during the different stressful virtual scenarios, subjects completed the Visual Analogue Scale for Anxiety (VAS-A) and the State-Trait Anxiety Inventory Form Y-1 (STAI-Y1) at the baseline and after each scenario. During stressful exposition, participant physiological parameters were also recorded. Besides the stressful scenarios (see Table 2), each participant was assessed in a neutral virtual environment and in one where they completed a cognitive task in front of a virtual audience. This allowed the therapist to identify the situations inducing the highest level of stress. At the end of this session, the clinician explained to the participant how to use the smartphone and the body-worn wireless sensor.

Training Session (Sessions 2-9)

The following sessions were dedicated to teaching participants how to cope with stress, through cognitive restructuring techniques and relaxation exercises. Each session lasted about 1 hour and was divided into four parts: homework checking,

exposure to a stressful VR environment, relaxation exercises (with or without biofeedback), and a homework assignment. The clinician decided with the participant the specific stress scenarios to work on in the exposure (cognitive restructuring) during at least two consecutive sessions. Cognitive restructuring was used to help patients identify and challenge their erroneous beliefs and interpretations. Specifically, patients were taught to look at their negative beliefs, look for possible alternative explanations and ways of thinking, and evaluate the pros and cons of maintaining them. Relaxation was induced through the immersion in a natural scenario selected by the subject, where they could move and interact. The experience also integrated different pre-recorded audio narratives describing the specific setting and guiding the execution of different progressive muscle relaxation/deep breathing exercises. The scenario was experienced with or without VR biofeedback, in alternate ways during sessions. At the end of the sessions, the clinician explained to the participant how to use the smartphone and the body-worn wireless sensor to check the level of contextual stress and do the contextualized homework.

Follow-up Session (Session 10)

To verify the efficacy of the training, during the final session participants were reassessed through the administration of psychometric questionnaires. Moreover, participants were re-exposed to the different stressful virtual scenarios, following the same procedure of the assessment session (Session 1). At the end of this assessment, the clinician discussed with the participant the protocol and its perceived efficacy/usefulness.

Control Group

The control group used a protocol based on traditional cognitive behavioral techniques, following the same structure (10 one-hour sessions in 5 weeks) and the same assessment points of the EG, but without the use of any technological tool. Exposure therapy designed for the control group involved imaginal exposure to the same stressful situations reproduced in virtual reality (see Table 2). Patients were instructed to close their eyes and experience the stressful situation by imagining that it was currently happening. An audio recording provided a detailed description of the context, the participants, the physical sensations, and emotional reactions.

Guided imagery was also used to teach relaxation exercises. Like in the EG, the imagery experience was supported by the same pre-recorded audio narratives describing the specific scenario and guiding the execution of different progressive muscle relaxation relaxation/deep breathing exercises.

CG participants did not use the mobile phone for stress assessment, but a traditional diary in which they recorded every relevant stressful event. Moreover, participants' homework was a self-help book about stress management. Topics included identifying and fully understanding what stress is, how stress affects our performance, the importance of becoming aware of stress, and simple strategies to make desired changes to reduce stress. See Multimedia Appendix 3 for the EG and CG study protocols.

Wait-List Group

The wait-list group did not receive treatment. They completed all facets of the pre-test and post-test (after 5 weeks), similar to those individuals included in the other groups.

Assessment

Questionnaires

To assess the effects of the stress management protocols, several questionnaires were used at different points of time.

Clinical Assessment

To exclude participants suffering from DSM-IV-TR Axis I disorders, the recruited sample was assessed before the start of the training using the MINI semistructured interview [45,46]. The MINI is a short diagnostic, structured interview that enables researchers to make diagnoses of psychiatric disorders according to DSM-IV or the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10). The administration time of the interview is approximately 15 minutes and was designed for epidemiological studies and multicentered clinical trials.

Primary Outcome Measures

Overview

The following questionnaires were administered offline under a therapist's supervision to each participant at pre-treatment, and upon completion of the trial (after 5 weeks).

State-Trait Anxiety Inventory Form

State-Trait Anxiety Inventory form (STAI) [72,73] consists of two scales containing 20 items each that measure anxiety in adults. The STAI clearly differentiates between the temporary condition of "state anxiety" (STAI Form Y-1, also known as STAI-Y1) and the more general and long-standing quality of "trait anxiety" (STAI Form Y-2, also known as STAI-Y2). For the initial and the final assessment in the trial, we used the STAI Y2, that is, the trait version of the STAI, which measures characteristic tendencies to be anxious.

Coping Orientation to the Problems Experienced—New Italian Version

Coping Orientation to the Problems Experienced (COPE) Inventory was originally developed to assess a broad range of coping responses [74]. COPE-NIV represents a useful and validated tool to assess different coping dimensions in an Italian context [75]. It consist of five scales, for a total of 60 items altogether: use of emotional support, denial, positive attitude, problem focused, and religious coping. This inventory can be used to assess trait coping (the usual way people cope with stress in everyday life) and state coping (the particular way people cope with a specific stressful situation).

Perceived Stress Scale

The Perceived Stress Scale (PSS) [76] is a 10-item self-reported measure designed to deal with the degree to which situations in an individual's life are appraised as stressful. It was originally developed as a 14-item scale that assessed the perception of stressful experiences over the previous month using a 5-point Likert scale. Later, the authors reported that the 10-item version

(PSS-10) showed stronger psychometric characteristics in comparison to the 14-item scale [77].

Psychological Stress Measure

The Psychological Stress Measure (PSM) [78,79] consists of 49 items based on the various individual perceptions of the cognitive, physiological, and behavioral state of subjects. PSM provides a global score of stress and some partial subscores. Patients are asked to answer on the basis of how they have been feeling in the last 4-5 days. The global score of the PSM is compared with ground truth scores, which give threshold cut-offs on the basis of the gender (103 for male and 110 for female subjects).

Satisfaction With Life Scale

The Satisfaction with Life Scale (SWLS) [80,81] is a measure of life satisfaction (subjective well-being). The 5-item questionnaire is designed to measure global cognitive judgments of satisfaction with one's own life. Qualitative feedback was also obtained using a semistructured questionnaire at pre-treatment, and upon completion of the trial.

Secondary Outcome Measures

Overview

Individuals were also assessed at the beginning and at the end of each of the 8 training sessions using the following questionnaires.

State-Trait Anxiety Inventory Form Y-1

STAI Y1 [72,73] addresses state anxiety, which could be defined as a temporary emotional condition characterized by apprehension, tension, and fear about a particular situation or activity. This inventory consists a of a 20-item scale, like the STAI Y2.

Visual Analogue Scale for Anxiety

Visual Analogue Scale for Anxiety (VAS-A) is an instrument that measures anxiety across a continuum. It is a horizontal line, 100 millimeters in length, anchored by word descriptors at each end (No anxiety; Very severe anxiety). Individuals mark on the line the point that they feel represents their perception of their own current state. The VAS-A score is determined by measuring in millimeters from the left end of the line to the point that the person marks.

Power Analysis

A power analysis was conducted to determine the appropriate number of participants needed for the current study. With an established alpha level of .05, 0.80 power, and a preliminary research based effect size of 0.272 (n=15, treatment vs non-treatment using the STAI Y2 score), a sample size of 30 participants for the group was enough to test the hypothesis of significant differences between groups. Using the trial sample size of 40 participants for groups, we achieved a power of more than 0.9.

Statistical Analysis

Our primary end points are the change of the level of chronic "trait" stress, perceived stress, and coping skills from baseline (pre) to the end of intervention (post). Secondary outcomes included the change in situational "state" stress scores from start (pre) to the end of intervention (post) in each treatment session.

For primary analyses, stress and coping scores were assessed by analysis of covariance (ANCOVA), with post-treatment scores as the baseline variable, while the pre-treatment scores served as covariates. This approach allowed us to *"adjust"* posttest scores for the variability on the pretest ones produced by the block design [82]. Several assumptions were met to use these techniques: (1) the use of interval or ratio data, (2) equally and normally distributed deviations in scores, (3) existence of a linear relationship between the dependent variable and the covariate (pretest scores), and (4) random assignment to groups.

A two-sided *P* value of .05 or less was considered to be statistically significant. For secondary analyses, the experimental and control groups were simultaneously taken into the analysis of variance model for repeated measures (T=8). Differential effects of the treatments were determined using post-hoc analyses. In particular, to reduce the risk of type I errors, we used the Bonferroni post-hoc procedure with an adjusted Experiment-wise Error Rate (EER): 0.05 for each variable in a three-group analysis and 0.025 for each variable in a four-group analysis [83]. Prior to analysis, the distributions for the outcome variables were examined. We detected univariate outliers using boxplots.

Results

Primary Outcome Variables

Outcome data were available for all the participants at the end of treatment (see Table 2). The one-way ANCOVA on the pre-post-treatment scores showed a significant group effect ($F_{2,107}=4.42$; $P<.014$; effect size=0.74) on the primary outcome variable of chronic "trait" anxiety (STAI Y2). While no significant differences were found in either the WL group or the CG between pre- and post-measurements, the EG was able to obtain a significant decrease in anxiety (1.2%). These data were confirmed by post-hoc analyses: they revealed significant differences between the EG and the CG ($P<.05$), and between the EG and the WL ($P<.01$).

We then used ANCOVA to analyze the pre-post changes in the level of coping skills. On one side, no significant differences were found in WL between the pre- and post-measurements. On the other, the two treatment groups were able to obtain significant coping skill improvement in four subscales (see Table 2) with a significant greater improvement for the EG ($P<.05$) in the Emotional Support skill. The analysis on the other stress scales (PSS, PSM) revealed a significant reduction in both treatments groups (see Table 3).

Post-hoc analyses did not reveal significant differences among the treatments, although the percentage of decrease in stress scales for the EG was higher (PSS: EG=21%, CG=16%; PSM: EG=13%, CG=5%). No differences were found in the WL group. Finally, we did not find difference between the three groups in the pre-post quality of life (SWLS) scores.

We also compared the two blocks (teachers and nurses) to check for differences. In general, nurses obtained slightly better values than teachers in most of the outcome variables for both treatments. We also found limited but significant differences in the STAI Y2 reduction (Teachers: -0.44+/-8.9; Nurses:

-3.2+/-5.3; P=.44, effect size=0.04), in the COPE-NIV Emotional Support skill increase (Teachers: -0.3+/-6.3; Nurses: 2.3+/-5.4; P=.20, effect size=0.05) and in the SWLS quality of life increase (Teachers: -0.6+/-5.58; Nurses: 1.8+/-3.3; P=.005, effect size=0.07).

Table 3. ANCOVA results.

Variables		Group			ANCOVA			
		Experimental	Control	Waiting List	F	df	P	η^2
STAI Y2								
	Time 1	43.6 (11.2)	42.1 (11.1)	39.8 (8.71)	4.42	2, 107	.014	.74
	Time 2	38.2 (8.09)	41.8 (10.8)	40.2 (8.82)				
COPE-NIV								
Use of emotional support								
	Time 1	27.8 (8.23)	29.9 (7.04)	29.4 (7.17)	17.2	2, 111	<.001	.237
	Time 2	31.7 (7.62)	29.5 (6.82)	28.8 (7.59)				
Positive attitude								
	Time 1	27.9 (5.54)	29.1 (4.69)	31.7 (4.08)	16.7	2, 109	<.001	.234
	Time 2	31.4 (5.16)	29.4 (5.86)	30.9 (4.01)				
Problem focused								
	Time 1	28.6 (5.96)	29.2 (7.36)	31.3 (3.78)	6.34	2, 109	.002	.104
	Time 2	31.7 (7.62)	29.5 (6.48)	29.6 (6.02)				
Religious coping								
	Time 1	19.8 (5.01)	21.7 (5.39)	23.8 (5.63)	2.81	2, 111	.064	.048
	Time 2	21.1 (6.65)	21.3 (5.63)	23.2 (6.1)				
Denial								
	Time 1	24.1 (5.91)	21.8 (4.69)	23.1 (3.69)	18.6	2, 116	.009	.082
	Time 2	22.3 (5.13)	20.9 (4.28)	24 (5.27)				
PSS								
	Time 1	21.1 (7.95)	18 (5.94)	19.5 (6.64)	10.1	2, 111	<.001	.155
	Time 2	16. 6 (4.75)	15 (5.55)	19.8 (7.24)				
PSM								
	Time 1	91.4 (25.8)	86.1 (19.5)	92.3 (26.1)	2.2	2, 111	.115	.39
	Time 2	79.2 (18.3)	81.6 (19.8)	92.4 (23.2)				
SWLS								
	Time 1	23.4 (6.22)	24.2 (7.18)	26.5 (6.95)	3.43	2, 111	.036	.058
	Time 2	25.5 (6.57)	24.5 (7.53)	23.2 (6.1)				

Secondary Outcome Variables

To better understand the differences between the EG and CG, we used a repeated measures analysis of variance (ANOVA) (T=8) to analyze the changes in the situational "state" stress scores (STAI Y1 and VAS-A) before and after each treatment.

The analysis revealed significant TIME ($F_{7,511}$=3.8; P<.01; effect size=0.05) and TIME x GROUP (Quadratic contrast: $F_{1,73}$=4.23; P<.05; effect size=0.05) effects in the VAS-A scores. First, both groups progressively increased the relaxation level

achieved in a session during the protocol. Second, the pattern of this increase is curvilinear (ie, is represented by a curve with one bend) with a difference between the two groups (see Figure 6). EG showed a marked (>15%) early increase in the relaxation level (from Session 3) and a further improvement (>20%) in the next sessions; CG instead showed an initial increase of the relaxation level (>10%) from the first session, with a marked (>15%) increase in the relaxation level only in the last sessions (7-8). The visible peaks experienced by EG in Figure 6 correspond to the introduction of a new stress scenario used by

the clinician for VR exposure (see description of the training sessions).

The analysis of STAI scores revealed a similar pattern but with a lack of significant values due to the low statistical power (0.23). We then compared the two blocks (teachers and nurses) to check for differences. Again, nurses obtained higher, but not significant, relaxation values than teachers in all the sessions.

Figure 6. Mean VAS-A reduction (pre-post) in the 8 treatment sessions for both treatment groups.
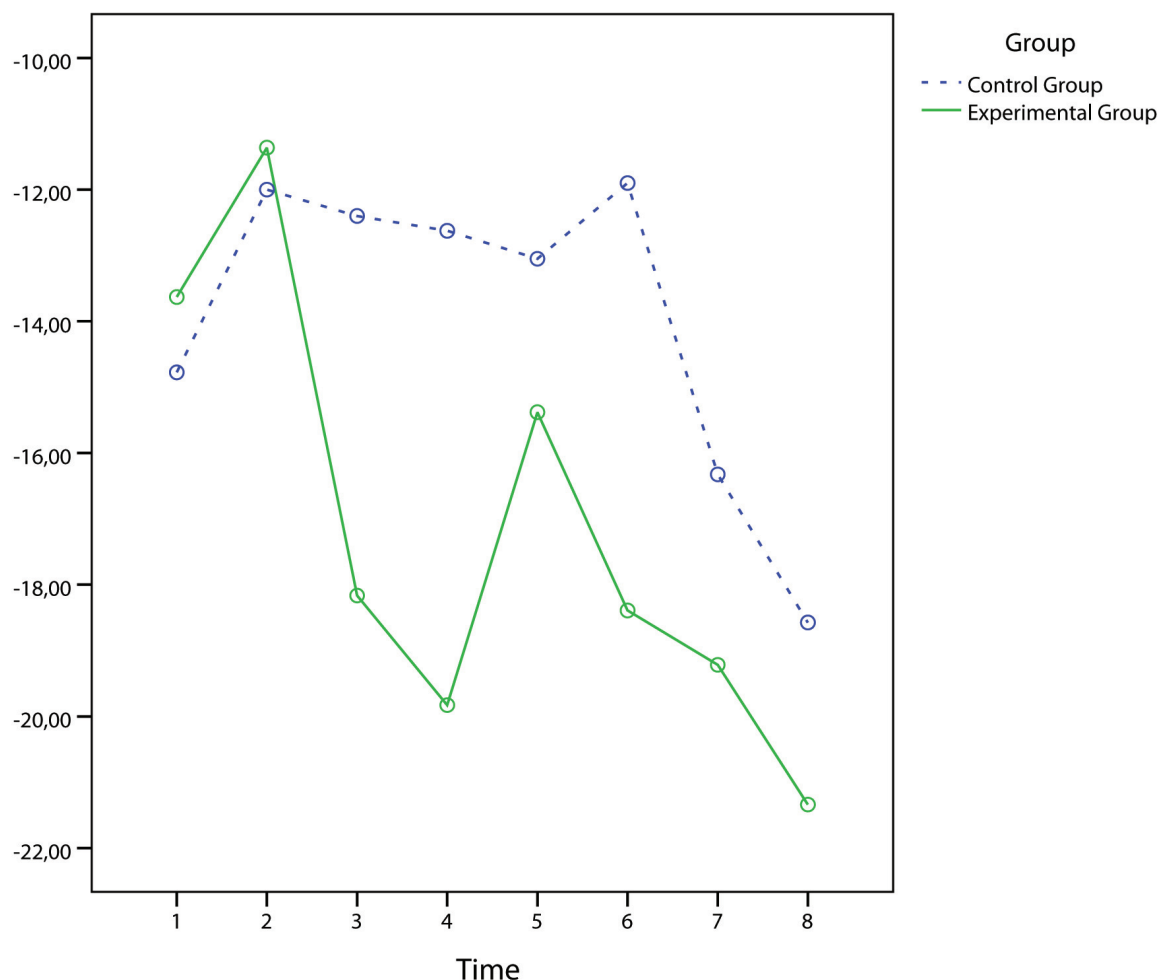
Discussion

Principal Results

Although both treatments (CBT and IR) were able to significantly reduce perceived stress (with a better outcome for IR), only participants who received IR reported a significant reduction (12% vs 0.5%) in chronic "trait" anxiety. This outcome is also better than that (11% reduction) obtained in a previous trial by a 5-week stress management meditation program where self-selected teachers were instructed to use meditation twice daily for 20-minute periods, both at school and at home [84]. In the IR protocol, no compulsory relaxation exercises were required outside the therapist's office, which is an obvious advantage for both unmotivated individuals and for individuals who are unable to understand the meditation technique or to apply it properly.

A similar pattern was found for coping skills. Both treatments were able to significantly increase most coping skills, but participants who received IR reported a significantly greater increase (14% vs 0.3%) in the Emotional Support skill than CBT. It is interesting to note that meditation/mindfulness stress management programs, unlike CBT ones, do not address coping skills in their protocol. However, different clinical trials showed long-term efficacy of coping skill training for better emotional control [85], quality of life [86,87], and resiliency enhancement [88].

In sum, the obtained data suggested both the clinical efficacy of IR and its enhanced efficacy over CBT and other existing protocols in the management of psychological stress.

Interreality as an Effective Clinical Protocol for Stress Management

Although different studies in the past may have evaluated the use of VR in the treatment of posttraumatic stress disorder, this is, to our knowledge, the first randomized controlled trial evaluating a technology-enhanced treatment program with active therapeutic involvement for the management of psychological stress.

The main issue in dealing with stress is that it is very personal. Thus, stress-related disorders cannot be explained simply on the basis of the adverse situations experienced by people. These disorders depend a great deal on how the person experiencing a stressor is put together psychologically and physically. So the focus for assessment, prediction, and treatment should be the situated experience of the person. And this is difficult to achieve using both CBT and the available VR protocols. The emotions and problems experienced by the patient in the real world are not directly assessed and/or addressed in real time by CBT; VR is a "closed" experience, separate from the emotions and behaviors experienced by the patient in the real world.

In this study, we provided initial evidence that a possible approach to overcome these issues is the use of the "Interreality" paradigm, which combines assessment and treatment within a hybrid environment, bridging physical and virtual worlds [40-42]. Specifically, the proposed protocol bridged experiential virtual scenarios (fully controlled by the therapist, used to learn coping skills and improve self-efficacy) with real-time monitoring and support (identifying critical situations and assessing clinical change), using advanced technologies (virtual worlds, wearable biosensors, and smartphones).

Possible explanations for the higher efficacy of the proposed approach are:

1. The added value offered to individuals by VR over guided imagery for the acquisition of behavioral and coping skills: As the findings indicate, VR is able to help individuals obtain a marked increase in the relaxation levels early in the therapy (from Session 3), while guided imagery is able to obtain the same relaxation levels only later on (from Session 7).

2. The added value offered to the individuals by real-time stress monitoring and support over traditional diary reporting: As suggested by qualitative reports, the IR sample strongly appreciated the possibility of having real-time stress monitoring and, if needed, use of the smartphone to experience the same relaxation protocols learned in the therapist's office in real-life settings.

3. The added value offered to the therapists by real-time stress monitoring over traditional diary reporting: In the early part of each session, IR therapists used the recorded real-time stress data to identify critical situations to be addressed in virtual scenarios. CBT therapists were not able to obtain similar detailed information from the diaries compiled by their subjects.

Strengths and Limitations

This study had several strengths as a direct test of the effects of an IR technology-enhanced protocol in the context of psychological stress management. The first strength of this study is that IR was compared with the best validated approach for stress management, CBT, as identified by the Cochrane Database of Systematic Reviews [36-38], demonstrating the added value offered by IR technology for psychological stress problems. Another important strength is the use, as primary outcome variables, of reliable theory-based measures of situational state and chronic trait anxiety. The third strength of IR is the use of VR to enhance the efficacy of biofeedback

intervention. Biofeedback training is regarded as a useful technique to reduce anxiety symptoms (eg, [89]). The most common limitation of biofeedback and relaxation training is that it requires time commitment and implementation effort on behalf of the patient, who can rely only on very simple audio and visual cues provided by the system to learn about body responses to stress. In VR-based biofeedback, elements of the virtual environment are directly modified by the patient's physiological parameters recorded in real time (eg, in the "Campfire" scenario, physiological parameters control the fire intensity, so that the reduction of the patient's physiological activation results in a reduction of the fire until it goes out). Thus, patients receives immediate feedback on their level of activation (as in the traditional biofeedback techniques), but with richer and more engaging 3D visual cues [90]. Unfortunately, our experimental design does not allow determination of the relative benefits of VR experiences and VR-based biofeedback intervention in reducing stress and anxiety, which is a future research goal.

However, the present findings should not be considered definitive. First, the study did not include a follow-up assessment of behavior maintenance in the long term. We have planned a follow-up research study to assess the participant samples again at 6-month and 12-month intervals.

A further limitation is that the study did not include any measure of physiological stress (eg, cortisol levels in blood, saliva, or urine samples). However, we justified this choice in the Methods following the results of a recent study demonstrating the low sensitivity of cortisol levels for individuals under long-term stress exposure [48].

Third, while our study design provided a strong test of the efficacy of the IR protocol, it did not allow an evaluation of the specific effectiveness of the different technological tools included in it. Further research is needed to identify the effective elements of the IR protocol and the optimal amount of technological intervention needed. As underlined clearly by qualitative reports, the usability, invasiveness, and complexity of the provided technology are the main barriers for a wide use of the proposed protocol: 2 out of 3 teachers and 1 out of 4 nurses involved in the EG evaluated as "high" or "very high" the amount of technological effort required by the protocol. In particular, the main issues indicated by the participants are the duration/charging of the batteries for both smartphone and biosensors, the Bluetooth pairing of the smartphone with the biosensors, the invasiveness of the biosensors, and the difficulty in consulting real-time stress data in some working situations (eg, teaching for teachers, consulting for nurses).

Finally, the IR technology–enhanced treatment is more expensive for both therapists (VR hardware including computer: €2200; biosensors: €2500) and patients (smartphone: €650; biosensor: €1000). Nevertheless, the rapid technological development is reducing the cost of the required technology. For example, at the end of the trial, the cost of a comparable body-worn wireless sensor decreased from €1000 (Empatica E3 Wrist Sensor) to €150 (Angel Health Monitor).

Conclusions

Our findings provide initial evidence that a technology-enhanced IR protocol provides better outcomes than the traditionally accepted gold standard for psychological stress treatment, CBT. However, a follow-up study assessing behavior maintenance in the long term is needed in order to provide additional evidence of the effectiveness of IR. In particular, unlike CBT, IR was able to significantly reduce chronic trait anxiety and better improve emotion-related coping skills. Consequently, these findings constitute a sound foundation and rationale for future research in technology-enhanced protocols for psychological stress management. In particular, our study provides the evidence required to justify carrying out much larger follow-up trials to identify the most effective technological interventions and the optimal amount of technological support needed.

Authors' Contributions

RG, GAnd, PF, RS, GAl, WB, and PG contributed to the design of the study. RG prepared the initial drafts of the paper. RG, GAnd, PF, SS, and WB critically discussed and reviewed the initial drafts of the paper. PG, BG, TG, and VC designed and developed some of the technologies used in the studies. PF, ML, SC, BMari, RS, Bmarg, CG, VN, and GAnn were involved in the treatment and in data collection CP, SS, PF, and RG carried out the analysis and interpretation of data. All authors contributed to the final draft of the paper and approved the final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH Checklist V1.6.2 [43].

[PDF File (Adobe PDF File), 2MB - jmir_v16i7e167_app1.pdf]

Multimedia Appendix 2

Stressful scenarios for the project.

[MP4 File (MP4 Video), 11MB - jmir_v16i7e167_app2.mp4]

Multimedia Appendix 3

EG and CG study protocols.

[PDF File (Adobe PDF File), 133KB - jmir_v16i7e167_app3.pdf]

References

1. Riva G. Ambient intelligence in health care. Cyberpsychol Behav 2003 Jun;6(3):295-300. [doi: 10.1089/109493103322011597] [Medline: 12855086]
2. Riva G. Virtual reality: an experiential tool for clinical psychology. British Journal of Guidance & Counselling 2009 Aug;37(3):337-345. [doi: 10.1080/03069880902957056]
3. Preziosa A, Grassi A, Gaggioli A, Riva G. Therapeutic applications of the mobile phone. British Journal of Guidance & Counselling 2009 Aug;37(3):313-325. [doi: 10.1080/03069880902957031]
4. Gorini A, Gaggioli A, Riva G. Virtual worlds, real healing. Science 2007 Dec 7;318(5856):1549. [doi: 10.1126/science.318.5856.1549b] [Medline: 18063769]
5. Gorini A, Gaggioli A, Vigna C, Riva G. A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology. J Med Internet Res 2008;10(3):e21 [FREE Full text] [doi: 10.2196/jmir.1029] [Medline: 18678557]
6. Wiederhold BK, Wiederhold MD. The future of cybertherapy: improved options with advanced technologies. Stud Health Technol Inform 2004;99:263-270. [Medline: 15295156]
7. Riva G, Molinari E, Vincelli F. Interaction and presence in the clinical relationship: virtual reality (VR) as communicative medium between patient and therapist. IEEE Trans Inf Technol Biomed 2002 Sep;6(3):198-205. [Medline: 12381035]
8. Riva G. Virtual reality in psychotherapy: review. Cyberpsychol Behav 2005 Jun;8(3):220-30; discussion 231. [doi: 10.1089/cpb.2005.8.220] [Medline: 15971972]
9. Moore K, Wiederhold BK, Wiederhold MD, Riva G. Panic and agoraphobia in a virtual world. Cyberpsychol Behav 2002 Jun;5(3):197-202. [Medline: 12123240]

XSL•FO
RenderX

10. Parsons TD, Rizzo AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. J Behav Ther Exp Psychiatry 2008 Sep;39(3):250-261. [doi: 10.1016/j.jbtep.2007.07.007] [Medline: 17720136]

11. Saunders T, Driskell JE, Johnston JH, Salas E. The effect of stress inoculation training on anxiety and performance. Journal of Occupational Health Psychology 1996;1(2):170-186. [doi: 10.1037/1076-8998.1.2.170]

12. Difede J, Hoffman H, Jaysinghe N. Innovative use of virtual reality technology in the treatment of PTSD in the aftermath of September 11. Psychiatr Serv 2002 Sep;53(9):1083-1085. [Medline: 12221305]

13. Difede J, Hoffman HG. Virtual reality exposure therapy for World Trade Center Post-traumatic Stress Disorder: a case report. Cyberpsychol Behav 2002 Dec;5(6):529-535. [Medline: 12556115]

14. Rothbaum BO. Using virtual reality to help our patients in the real world. Depress Anxiety 2009;26(3):209-211. [doi: 10.1002/da.20556] [Medline: 19253416]

15. Wood DP, Murphy J, McLay R, Koffman R, Spira J, Obrecht RE, et al. Cost effectiveness of virtual reality graded exposure therapy with physiological monitoring for the treatment of combat related post traumatic stress disorder. Stud Health Technol Inform 2009;144:223-229. [Medline: 19592768]

16. Wood DP, Murphy JA, Center KB, Russ C, McLay RN, Reeves D, et al. Combat related post traumatic stress disorder: a multiple case report using virtual reality graded exposure therapy with physiological monitoring. Stud Health Technol Inform 2008;132:556-561. [Medline: 18391367]

17. Rothbaum BO, Hodges L, Alarcon R, Ready D, Shahar F, Graap K, et al. Virtual reality exposure therapy for PTSD Vietnam Veterans: a case study. J Trauma Stress 1999 Apr;12(2):263-271. [doi: 10.1023/A:1024772308758] [Medline: 10378165]

18. Rothbaum BO, Hodges LF, Ready D, Graap K, Alarcon RD. Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. J Clin Psychiatry 2001 Aug;62(8):617-622. [Medline: 11561934]

19. Wood DP, Murphy J, Center K, McLay R, Reeves D, Pyne J, et al. Combat-related post-traumatic stress disorder: a case report using virtual reality exposure therapy with physiological monitoring. Cyberpsychol Behav 2007 Apr;10(2):309-315. [doi: 10.1089/cpb.2006.9951] [Medline: 17474853]

20. Gamito P, Oliveira J, Morais D, Oliveira S, Duarte N, Saraiva T, et al. Virtual Reality Therapy Controlled Study for War Veterans with PTSD. Preliminary Results. Stud Health Technol Inform 2009;144:269-272. [Medline: 19592779]

21. Gerardi M, Rothbaum BO, Ressler K, Heekin M, Rizzo A. Virtual reality exposure therapy using a virtual Iraq: case report. J Trauma Stress 2008 Apr;21(2):209-213 [FREE Full text] [doi: 10.1002/jts.20331] [Medline: 18404648]

22. Walshe DG, Lewis EJ, Kim SI, O'Sullivan K, Wiederhold BK. Exploring the use of computer games and virtual reality in exposure therapy for fear of driving following a motor vehicle accident. Cyberpsychol Behav 2003 Jun;6(3):329-334. [doi: 10.1089/109493103322011641] [Medline: 12855091]

23. McLay RN, McBrien C, Wiederhold MD, Wiederhold BK. Exposure therapy with and without virtual reality to treat PTSD while in the combat theater: a parallel case series. Cyberpsychol Behav Soc Netw 2010 Feb;13(1):37-42. [Medline: 20528291]

24. Rizzo AA, Difede J, Rothbaum BO, Johnston S, McLay RN, Reger G, et al. VR PTSD exposure therapy results with active duty OIF/OEF combatants. Stud Health Technol Inform 2009;142:277-282. [Medline: 19377167]

25. Difede J, Cukor J, Patt I, Giosan C, Hoffman H. The application of virtual reality to the treatment of PTSD following the WTC attack. Ann N Y Acad Sci 2006 Jul;1071:500-501. [doi: 10.1196/annals.1364.052] [Medline: 16891607]

26. McLay RN, Wood DP, Webb-Murphy JA, Spira JL, Wiederhold MD, Pyne JM, et al. A randomized, controlled trial of virtual reality-graded exposure therapy for post-traumatic stress disorder in active duty service members with combat-related post-traumatic stress disorder. Cyberpsychol Behav Soc Netw 2011 Apr;14(4):223-229. [doi: 10.1089/cyber.2011.0003] [Medline: 21332375]

27. Astin MC, Lawrence KJ, Foy DW. Posttraumatic stress disorder among battered women: risk and resiliency factors. Violence Vict 1993;8(1):17-28. [Medline: 8292561]

28. Bryant RA, Harvey AG. Initial posttraumatic stress responses following motor vehicle accidents. J Trauma Stress 1996 Apr;9(2):223-234. [Medline: 8731544]

29. Shekhar A, Truitt W, Rainnie D, Sajdyk T. Role of stress, corticotrophin releasing factor (CRF) and amygdala plasticity in chronic anxiety. Stress 2005 Dec;8(4):209-219. [doi: 10.1080/10253890500504557] [Medline: 16423710]

30. Laposa JM, Alden LE, Fullerton LM. Work stress and posttraumatic stress disorder in ED nurses/personnel. J Emerg Nurs 2003 Feb;29(1):23-28. [Medline: 12556825]

31. MacDonald HA, Colotla V, Flamer S, Karlinsky H. Posttraumatic stress disorder (PTSD) in the workplace: a descriptive study of workers experiencing PTSD resulting from work injury. J Occup Rehabil 2003 Jun;13(2):63-77. [Medline: 12708101]

32. 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. Journal of Environmental and Public Health 2012;2012:1-21 [FREE Full text] [doi: 10.1155/2012/515874]

33. Lamontagne AD, Keegel T, Louie AM, Ostry A, Landsbergis PA. A systematic review of the job-stress intervention evaluation literature, 1990-2005. Int J Occup Environ Health 2007;13(3):268-280. [Medline: 17915541]

34. Stetz MC, Kaloi-Chen JY, Turner DD, Bouchard S, Riva G, Wiederhold BK. The effectiveness of technology-enhanced relaxation techniques for military medical warriors. Mil Med 2011 Sep;176(9):1065-1070. [Medline: 21987967]

35. Cohen S, Janicki-Deverts D, Miller GE. Psychological stress and disease. JAMA 2007 Oct 10;298(14):1685-1687. [doi: 10.1001/jama.298.14.1685] [Medline: 17925521]

36. Bisson J, Andrew M. Psychological treatment of post-traumatic stress disorder (PTSD). Cochrane Database Syst Rev 2005(2):CD003388. [doi: 10.1002/14651858.CD003388.pub2] [Medline: 15846661]

37. Whalley B, Rees K, Davies P, Bennett P, Ebrahim S, Liu Z, et al. Psychological interventions for coronary heart disease. Cochrane Database Syst Rev 2011(8):CD002902. [doi: 10.1002/14651858.CD002902.pub3] [Medline: 21833943]

38. Thomson AB, Page L. Psychotherapies for hypochondriasis. Cochrane Database Syst Rev 2007(4):CD006520. [doi: 10.1002/14651858.CD006520.pub2] [Medline: 17943915]

39. Riva G, Raspelli S, Algeri D, Pallavicini F, Gorini A, Wiederhold BK, et al. Interreality in practice: bridging virtual and real worlds in the treatment of posttraumatic stress disorders. Cyberpsychol Behav Soc Netw 2010 Feb;13(1):55-65. [Medline: 20528294]

40. Riva G. Interreality: A New Paradigm for E-health. Stud Health Technol Inform 2009;144:3-7. [Medline: 19592718]

41. Pallavicini F, Gaggioli A, Raspelli S, Cipresso P, Serino S, Vigna C, et al. Interreality for the management and training of psychological stress: study protocol for a randomized controlled trial. Trials 2013;14:191 [FREE Full text] [doi: 10.1186/1745-6215-14-191] [Medline: 23806013]

42. Riva G, Raspelli S, Pallavicini F, Grassi A, Algeri D, Wiederhold BK, et al. Interreality in the management of psychological stress: a clinical scenario. Stud Health Technol Inform 2010;154:20-25. [Medline: 20543263]

43. 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]

44. Sunnerhagen KS, Danielsson A, Rafsten L, Björkdahl A, Axelsson Å, Nordin Å, et al. Gothenburg very early supported discharge study (GOTVED) NCT01622205: a block randomized trial with superiority design of very early supported discharge for patients with stroke. BMC Neurol 2013;13:66 [FREE Full text] [doi: 10.1186/1471-2377-13-66] [Medline: 23800106]

45. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]

46. Rossi A, Alberio R, Porta A, Sandri M, Tansella M, Amaddeo F. The reliability of the Mini-International Neuropsychiatric Interview--Italian version. J Clin Psychopharmacol 2004 Oct;24(5):561-563. [Medline: 15349020]

47. Barth J. Neglecting the subjective relevance of stress on health: a protective strategy for your heart? Eur Heart J 2013 Sep;34(34):2658-2659. [doi: 10.1093/eurheartj/eht289] [Medline: 23886916]

48. Faresjö Å, Theodorsson E, Chatziarzenis M, Sapouna V, Claesson HP, Koppner J, et al. Higher perceived stress but lower cortisol levels found among young Greek adults living in a stressful social environment in comparison with Swedish young adults. PLoS One 2013;8(9):e73828 [FREE Full text] [doi: 10.1371/journal.pone.0073828] [Medline: 24066077]

49. Kaluza G. Changing unbalanced coping profiles-A prospective controlled intervention trial in worksite health promotion. Psychology & Health 2000 May;15(3):423-433. [doi: 10.1080/08870440008402003]

50. Meichenbaum D. Stress Inoculation training. New York: Pergamon Press; 1985:1-115.

51. Kaluza G. Evaluation von Stressbewältigungstrainings in der primären Prävention – eine Metaanalyse (quasi-)experimenteller Feldstudien (Evaluation of stress management trainings in primary prevention – a meta-analysis of (quasi-)experimental field studies). Zeitschrift für Gesundheitspsychologie 1997;5:149-169 [FREE Full text]

52. Meichenbaum D. Cognitive-behavior modification: an integrative approach. New York: Plenum Press; 1977.

53. Stetz MC, Thomas ML, Russo MB, Stetz TA, Wildzunas RM, McDonald JJ, et al. Stress, mental health, and cognition: a brief review of relationships and countermeasures. Aviat Space Environ Med 2007 May;78(5 Suppl):B252-B260. [Medline: 17547326]

54. Foley FW, Bedell JR, LaRocca NG, Scheinberg LC, Reznikoff M. Efficacy of stress-inoculation training in coping with multiple sclerosis. J Consult Clin Psychol 1987 Dec;55(6):919-922. [Medline: 3693660]

55. Moore K, Altmaier EM. Stress inoculation training with cancer patients. Cancer Nurs 1981 Oct;4(5):389-393. [Medline: 6913430]

56. Ross MJ, Berger RS. Effects of stress inoculation training on athletes' postsurgical pain and rehabilitation after orthopedic injury. J Consult Clin Psychol 1996 Apr;64(2):406-410. [Medline: 8871425]

57. Cecil MA, Forman SG. Effects of stress inoculation training and coworker support groups on teachers' stress. Journal of School Psychology 1990 Jun;28(2):105-118. [doi: 10.1016/0022-4405(90)90002-O]

58. von Baeyer C, Krause L. Effectiveness of stress management training for nurses working in a burn treatment unit. Int J Psychiatry Med 1983;13(2):113-126. [Medline: 6642873]

59. Mace R, Carroll D. Stress inoculation training to control anxiety in sport: two case studies in squash. Br J Sports Med 1986 Sep;20(3):115-117 [FREE Full text] [Medline: 3779338]

60. Riva G, Carelli L, Gaggioli A, Gorini A, Vigna C, Corsi R, et al. NeuroVR 1.5 - a free virtual reality platform for the assessment and treatment in clinical psychology and neuroscience. Stud Health Technol Inform 2009;142:268-270. [Medline: 19377165]

61. Riva G, Gaggioli A, Grassi A, Raspelli S, Cipresso P, Pallavicini F, et al. NeuroVR 2--a free virtual reality platform for the assessment and treatment in behavioral health care. Stud Health Technol Inform 2011;163:493-495. [Medline: 21335845]

62. Riva G, Gaggioli A, Villani D, Preziosa A, Morganti F, Corsi R, et al. NeuroVR: an open source virtual reality platform for clinical psychology and behavioral neurosciences. Stud Health Technol Inform 2007;125:394-399. [Medline: 17377310]

63. Gorini A, Riva G. The potential of virtual reality as anxiety management tool: a randomized controlled study in a sample of patients affected by generalized anxiety disorder. Trials 2008;9:25 [FREE Full text] [doi: 10.1186/1745-6215-9-25] [Medline: 18457580]

64. Gorini A, Pallavicini F, Algeri D, Repetto C, Gaggioli A, Riva G. Virtual reality in the treatment of generalized anxiety disorders. Stud Health Technol Inform 2010;154:39-43. [Medline: 20543266]

65. Manzoni GM, Pagnini F, Castelnuovo G, Molinari E. Relaxation training for anxiety: a ten-years systematic review with meta-analysis. BMC Psychiatry 2008;8:41 [FREE Full text] [doi: 10.1186/1471-244X-8-41] [Medline: 18518981]

66. Stetz MC, Bouchard S, Wiederhold BK, Riva G, Folen RA. The receptiveness of stress management techniques by military personnel. Stud Health Technol Inform 2009;144:125-127. [Medline: 19592747]

67. Ferrer-García M, Gutiérrez-Maldonado J, Caqueo-Urízar A, Moreno E. The validity of virtual environments for eliciting emotional responses in patients with eating disorders and in controls. Behav Modif 2009 Nov;33(6):830-854. [doi: 10.1177/0145445509348056] [Medline: 19822737]

68. Manzoni GM, Preziosa A, Pagnini F, Castelnuovo G, Molinari E, Riva G. New Technologies and Relaxation: An exploratory study on obese patients with emotional eating. The Role Of Presence. Journal of Cybertherapy and Rehabilitation 2008;1(2):182-192.

69. Pallavicini F, Algeri D, Repetto C, Gorini A, Riva G. Biofeedback, virtual reality and mobile phones in the treatment of Generalized Anxiety Disorder (GAD): A phase-2 controlled clinical trial. Journal of Cybertherapy and Rehabiliation 2009;2(4):315-327.

70. Gaggioli A, Pioggia G, Tartarisco G, Baldus G, Ferro M, Cipresso P, et al. A system for automatic detection of momentary stress in naturalistic settings. Stud Health Technol Inform 2012;181:182-186. [Medline: 22954852]

71. Tartarisco G, Baldus G, Corda D, Raso R, Arnao A, Ferro M, et al. Personal Health System architecture for stress monitoring and support to clinical decisions. Computer Communications 2012 Jun;35(11):1296-1305. [doi: 10.1016/j.comcom.2011.11.015]

72. Spielberger CD, Gorsuch RL, Lushene RE. Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press; 1970.

73. Macor A, Pedrabissi L, Santinello M. Ansia di stato e di tratto: ulteriore contributo alla verifica della validita psicometrica e teorica dello S. T.A.I. forma Y di Spielberger. Psicologia e società 1990;15(1/3):67-74.

74. Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: a theoretically based approach. J Pers Soc Psychol 1989 Feb;56(2):267-283. [Medline: 2926629]

75. Sica C, Novara C, Dorz S, Sanavio E. Coping strategies: Evidence for cross-cultural differences? A preliminary study with the Italian version of coping orientations to problems experienced (COPE). Personality and Individual Differences 1997 Dec;23(6):1025-1029. [doi: 10.1016/S0191-8869(97)00112-8]

76. Cohen S. Contrasting the hassle scale and the perceived stress scale. American Psychologist 1986;41:717-718.

77. Spacapan S, Oskamp S. Perceived stress in a probability sample of the United States. In: The Social psychology of health. Newbury Park, Calif: Sage Publications; 1988.

78. Lemyre L, Tessier R. Mesure de Stress Psychologique (MSP): Se sentir stressé-e. Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement 1988;20(3):302-321. [doi: 10.1037/h0079945]

79. Di Nuovo S, Rispoli L, Genta E. Misurare lo stress. Milano: Franco Angeli; 2000.

80. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. Journal of Personality Assessment 1985;49(1):71-75 [FREE Full text] [doi: 10.1207/s15327752jpa4901_13]

81. Prezza M, Costantini S. Sense of community and life satisfaction: investigation in three different territorial contexts. J. Community. Appl. Soc. Psychol 1998 May;8(3):181-194. [doi: 10.1002/(SICI)1099-1298(199805/06)8:3<181::AID-CASP436>3.0.CO;2-4]

82. Dimitrov DM, Rumrill PD. Pretest-posttest designs and measurement of change. Work 2003;20(2):159-165. [Medline: 12671209]

83. Stevens JP. Applied multivariate statistics for the social sciences. Mahwah, NJ: Lawrence Erlbaum Associates; 2002.

84. Anderson VL, Levinson EM, Barker W, Kiewra KR. The Effects of Meditation on Teacher Perceived Occupational Stress, State and Trait Anxiety, and Burnout. School Psychology Quarterly 1999;14(1):3-25.

85. Emery CF, Keefe FJ, France CR, Affleck G, Waters S, Fondow MD, et al. Effects of a brief coping skills training intervention on nociceptive flexion reflex threshold in patients having osteoarthritic knee pain: a preliminary laboratory study of sex differences. J Pain Symptom Manage 2006 Mar;31(3):262-269. [doi: 10.1016/j.jpainsymman.2005.07.008] [Medline: 16563320]

86. Grey M, Boland EA, Davidson M, Li J, Tamborlane WV. Coping skills training for youth with diabetes mellitus has long-lasting effects on metabolic control and quality of life. J Pediatr 2000 Jul;137(1):107-113. [doi: 10.1067/mpd.2000.106568] [Medline: 10891831]

87.   Grey M, Boland EA, Davidson M, Yu C, Tamborlane WV. Coping skills training for youths with diabetes on intensive therapy. Appl Nurs Res 1999 Feb;12(1):3-12. [Medline: 10048236]

88.   Jafari E, Eskandari H, Sohrabi F, Delavar A, Heshmati R. Effectiveness of coping skills training in relapse prevention and resiliency enhancement in people with substance dependency. Procedia - Social and Behavioral Sciences 2010 Jan;5:1376-1380. [doi: 10.1016/j.sbspro.2010.07.291]

89.   Ratanasiripong P, Ratanasiripong N, Kathalae D. Biofeedback Intervention for Stress and Anxiety among Nursing Students: A Randomized Controlled Trial. ISRN Nurs 2012;2012:827972 [FREE Full text] [doi: 10.5402/2012/827972] [Medline: 22811932]

90.   Repetto C, Gorini A, Vigna C, Algeri D, Pallavicini F, Riva G. The use of Biofeedback in Clinical Virtual Reality: The INTREPID Project. Journal of Visualized Experiments 2009;33:1-4 [FREE Full text]

Abbreviations

ANCOVA: analysis of covariance
CBT: cognitive behavioral therapy
CG: control group
COPE: Coping Orientation to Problems Experienced inventory
COPE-NIV: Coping Orientation to the Problems Experienced—New Italian Version
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th edition
EG: experimental group
IR: Interreality
MINI: Mini-International Neuropsychiatric Interview
PSM: Psychological Stress Measure
PSS: Perceived Stress Scale
PTSD: posttraumatic stress disorders
STAI: State-Trait Anxiety Inventory
SWLS: Satisfaction With Life Scale
WL: wait-list group
VAS-A: Visual Analogue Scale for Anxiety
VR: virtual reality

XSL•FO
RenderX

Original Paper

Comparative Analysis of Online Health Queries Originating From Personal Computers and Smart Devices on a Consumer Health Information Portal

Ashutosh Jadhav[1], MS; Donna Andrews[2]; Alexander Fiksdal[2], MA; Ashok Kumbamu[2], PhD; Jennifer B McCormick[2], PhD; Andrew Misitano[2]; Laurie Nelsen[2]; Euijung Ryu[2], PhD; Amit Sheth[1], PhD; Stephen Wu[2], PhD; Jyotishman Pathak[3], PhD

[1]Knoesis Ceneter, Computer Science and Engineering, Wright State University, Dayton, OH, United States

[2]Mayo Clinic, Rochester, MN, United States

[3]Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, United States

Corresponding Author:
Jyotishman Pathak, PhD
Division of Biomedical Statistics and Informatics
Department of Health Sciences Research
Mayo Clinic
200 First Street SW
Rochester, MN, 55905
United States
Phone: 1 (507) 538 8384
Fax: 1 (507) 284 0360
Email: pathak.jyotishman@mayo.edu

Abstract

Background: The number of people using the Internet and mobile/smart devices for health information seeking is increasing rapidly. Although the user experience for online health information seeking varies with the device used, for example, smart devices (SDs) like smartphones/tablets versus personal computers (PCs) like desktops/laptops, very few studies have investigated how online health information seeking behavior (OHISB) may differ by device.

Objective: The objective of this study is to examine differences in OHISB between PCs and SDs through a comparative analysis of large-scale health search queries submitted through Web search engines from both types of devices.

Methods: Using the Web analytics tool, IBM NetInsight OnDemand, and based on the type of devices used (PCs or SDs), we obtained the most frequent health search queries between June 2011 and May 2013 that were submitted on Web search engines and directed users to the Mayo Clinic's consumer health information website. We performed analyses on "Queries with considering repetition counts (QwR)" and "Queries without considering repetition counts (QwoR)". The dataset contains (1) 2.74 million and 3.94 million QwoR, respectively for PCs and SDs, and (2) more than 100 million QwR for both PCs and SDs. We analyzed structural properties of the queries (length of the search queries, usage of query operators and special characters in health queries), types of search queries (keyword-based, wh-questions, yes/no questions), categorization of the queries based on health categories and information mentioned in the queries (gender, age-groups, temporal references), misspellings in the health queries, and the linguistic structure of the health queries.

Results: Query strings used for health information searching via PCs and SDs differ by almost 50%. The most searched health categories are "Symptoms" (1 in 3 search queries), "Causes", and "Treatments & Drugs". The distribution of search queries for different health categories differs with the device used for the search. Health queries tend to be longer and more specific than general search queries. Health queries from SDs are longer and have slightly fewer spelling mistakes than those from PCs. Users specify words related to women and children more often than that of men and any other age group. Most of the health queries are formulated using keywords; the second-most common are wh- and yes/no questions. Users ask more health questions using SDs than PCs. Almost all health queries have at least one noun and health queries from SDs are more descriptive than those from PCs.

Conclusions: This study is a large-scale comparative analysis of health search queries to understand the effects of device type (PCs vs SDs) used on OHISB. The study indicates that the device used for online health information search plays an important role in shaping how health information searches by consumers and patients are executed.

Introduction

Background

Limited health literacy is associated with higher rates of hospitalizations, poorer health, and higher mortality [1,2]. Online health information plays a vital role in improving health literacy and helps online health information seekers (OHIS) make more informed health decisions. Over the past decade, Internet literacy and the number of Internet users have increased significantly [3-5]. With the growing availability of eHealth resources [6,7], consumers are increasingly using the Internet to seek health-related information. According to a 2013 Pew Survey [4], one in three American adults has gone online to find information about a specific medical condition. With the recent exponential increase in usage of smart devices (SDs), like smartphones or tablets, the percentage of people using smart devices to search for health information is also growing rapidly [8,9].

While there is some evidence [10] that the experience of online information searching varies depending on the device used (eg, smart devices vs personal computers or laptops [PCs]), little is known about how device choice impacts the structure of health information search queries generated by users. Understanding the effects of the device used (SDs vs PCs) for health information search would help us acquire more insights into online health information seeking behavior (OHISB). Such knowledge can be applied to improve the search experience and to develop more advanced next-generation knowledge and content delivery systems.

One of the most common ways to seek online health Information is via Web search engines such as Google, Yahoo!, etc [4]. According to the Pew Survey [4], approximately 8 in 10 online health inquiries initiate from a search engine. A typical online health information search process starts with the formulation of a health search query based on an OHIS information need. This query is typically submitted to a Web search engine, which subsequently leads to visiting one or more websites recommended by the search engine. In this paper, we study the effect of the devices used for health information search, concentrating on what information users search for and how health search queries are formulated.

Using the Mayo Clinic website's Web analytics tool (IBM NetInsight OnDemand [11]) and based on the type of devices used (PCs or SDs), we obtained the most frequent health search queries submitted from Web search engines that direct traffic to the Mayo Clinic webpages [12]. We selected search queries that are in the English language and collected between June

2011 and May 2013. We analyzed structural properties, types (keywords, wh-question, yes/no-questions), misspellings, and the linguistic structure of the health queries. We further categorized them based on health categories and demographic information mentioned (gender, age group, etc) in the queries. Our analysis suggests that the device used for online health information searching plays a significant role, altering the OHISB.

Significance of Current Study

Many previous studies have investigated OHISB. Researchers have used several approaches to understand OHISB including (1) focus groups and user surveys [13-20] and (2) analyzing health-related Web search query logs [21-32]. In the studies that involved focus groups and user surveys, researchers have analyzed characteristics associated with OHISB such as how people use the Internet for health information searching, their demographic information (age, gender, education level, etc), devices/Web search engines used for searching, OHISB in specific health conditions, and age groups [13-20]. Although these studies provide important insights into OHISB, their main limitation was the inclusion of a small number of participants (ranging from 100-2000 people). A second approach to studying OHISB is analyzing Web search logs from the health domain. Several previous studies have analyzed health search query logs with diverse objectives, such as health/epidemic surveillance [33-39], PubMed usage [40,41], and OHISB [21-32]. The studies focusing on OHISB [21-32] have studied a variety of aspects of health query logs, such as query length, health categories, relationship between OHISB and health care utilization [24], changes in health behavior with type of disease [21], and changes in OHISB with disease escalation from symptoms to serious illness [22,23].

Although the user experience for online health information searching varies with the device used (PCs/SDs) [10], there is a dearth of work relating OHISB with the device used for searching. In this study, we address this problem by analyzing large-scale health queries for both PCs and SDs to understand the effects of device type (PCs vs SDs) used for online health information seeking. Previous studies in generic search query log analysis have determined the importance of understanding linguistic structure of search queries as it has implications on information retrieval using Web search engines [42,43]. One of the contributions of this study is a comparative analysis of linguistic structure of health search queries from PCs and SDs. This study provides useful and interesting findings that can be leveraged in multiple ways. Some of the potential beneficiaries are (1) Web search engines: to understand health search query structure and complexity, and the occurrence of popular health

categories for PCs and SDs to improve query performance and accuracy for health information retrieval systems, (2) Websites that provide health information: to better understand online health information seekers' health information need, and better organize health information content for PCs and SDs users, (3) Health care providers: to better understand their patients and their health information interests, (4) Health care-centric application developers: to better understand OHISB for PCs and SDs and build applications around consumer's health information needs and priorities, and (5) online health information seekers: we anticipate that this work will help empower online health information seekers in their quest for health information and facilitate their health information search efforts by enabling the development of smarter and more sophisticated consumer health information delivery mechanisms.

Methods

Data Source

In this study, we collected health search queries originating from Web search engines (such as Google and Bing) that direct OHISs to the Mayo Clinic's consumer health information website [12], which is one of the top online health information website within the United States. The Mayo Clinic provides up-to-date, high-quality online health information produced by professional writers and editors. Our recent Web analytics statistics indicate that the Mayo Clinic website is visited by millions of unique visitors on average every day, and around 90% of the incoming traffic originated from Web search engines. The Mayo Clinic website is identical in terms of appearance and functionality for both PCs and SDs using standard Web search engines and Web browsers. This consistency as well as significant traffic to the website provide us with an excellent platform to conduct our study.

Dataset Creation

The Mayo Clinic website's Web analytics tool, IBM NetInsight OnDemand [11], keeps detailed information about incoming Web traffic from Web search engines to the Mayo Clinic website. The tool maintains information such as input search query (the original query from a Web search engine that brings an OHIS to the Mayo Clinic website), number of query repetitions (how many times the query has been searched within specified time period), and the visitor's Operating System (OS). PCs generally use Windows (98, 2000, Xp, Vista, 7, 8), Mac OS X, or Linux (such as Ubuntu and Redhat) operating systems while SDs use iOS (iPhone's OS), Android, Windows Mobile, and RIM BlackBerry operating systems. Since the Web analytics tool tracks information related to each user's OS type and individual searches, we are able to differentiate search queries by device type (PCs/SDs).

Using the Web analytics tool, we obtained one data report for each of the most frequent one million (based on the number of query repetitions) anonymized distinct queries in the English language launched from PCs and SDs for each month between June 2011 and May 2013 (24 months), totalling 48 data reports. Each search query appears uniquely in each data report and has an associated number of query repetitions. For each device type

(PCs and SDs), we aggregated 24 reports to create a single report with distinct queries. The dataset for PCs has 2.74 million queries, and the dataset for SDs has 3.94 million queries. While aggregating the search queries for PCs and SDs, we combined the repetition counts for each repeated query; for example, if a "diabetes" query has 5 repetitions in 1 month and 10 repetitions in another month, then the total number of repetition for the "diabetes" query is 15. Note that selecting the top queries for 2 years would be an easier approach for dataset creation, but in our case the data reports were available by month, thus we have to aggregate the data for each month to create the final analysis dataset.

Data Analysis

Overview

In this study, we performed analyses on "queries *with considering* repetition counts (QwR)" and "queries *without considering* repetition counts (QwoR)". Because the analysis performed with only QwR may overrepresent certain queries due to their large number of repetitions, we performed the analysis for both QwoR and QwR. The QwoR count is the same as the number of queries in the dataset. Hence for PCs, we have 2.74 million QwoR, and for SDs we have 3.94 million QwoR. We obtained the QwR count by aggregating number of repetitions for all the queries in the dataset. For both PCs and SDs, we got more than 100 million QwR. Due to Mayo Clinic's confidentiality policy, we are not able to disclose the exact number of QwR. We are reporting percentages of PC and SD queries.

Top Health Queries

The top search queries are the most commonly searched queries. To analyze the top health queries launched from PCs and SDs, we selected the top 100 search queries, from PCs and SDs, based on the descending order of number of query repetitions in the analysis dataset.

Health Categories

To analyze popular health categories that OHISs search for from PCs and SDs, we selected the following 8 health categories corresponding to the organization of health topics on popular health websites (Mayo Clinic, MedlinePlus [44], WebMD [45]): Symptoms, Causes, Complications, Tests and Diagnosis, Treatments and Drugs, Risk Factors, Prevention, Coping and Support. For example, Figure 1 shows different health categories for diabetes on the Mayo Clinic website, where each health category has a separate webpage with detailed information (browsable via navigating the left panel). Based on the semantics of an OHIS's input search query and a Web search engine's recommendations, users may land on one of the health category pages on the Mayo Clinic website. For this study, we aggregated all the incoming health search queries between June 2011 and May 2013 that land on a particular health category webpages. For example, we aggregated all the search queries that land on the "Symptoms" webpage for all the diseases and health conditions on the Mayo Clinic website. We analyzed the type of device (PC or SD) used for searches and the number of search queries to each health category.

Figure 1. Screenshot of Mayo Clinic website for Diabetes (left-side box highlights organization of health information based on health categories).
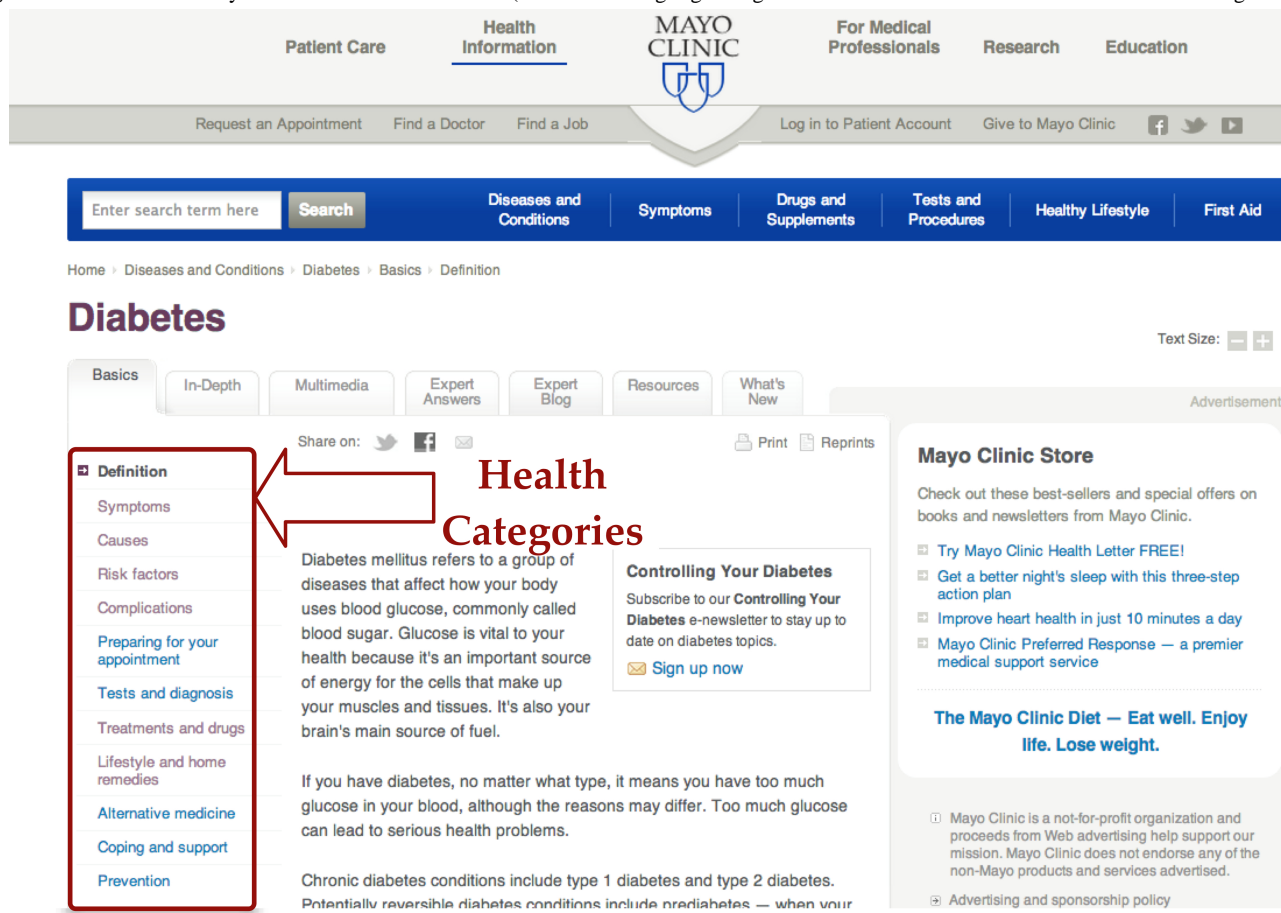
Categorization Based on the Information Mentioned in the Health Queries

In order to understand how often an OHIS mentions gender, age groups, and temporal references in the search queries, we categorized health queries using a dictionary-based approach. For each group, we created a lexicon by going through online English dictionaries [46-48] and a manual evaluation of words. For example, in the "Gender" group we considered Men (Man, men, male, boy, gent, gentleman, gentlemen) and Women (Woman, women, female, girl, ladies, lady). We also considered keywords' lexical variants; for example, boy, boys, etc. We categorized search queries from PCs and SDs by utilizing the lexicon for each category.

Health Query Length

To study the difference in health search query length for queries from PCs and SDs, we calculated search query length by computing the number of words (separated by white space) and the number of characters (excluding white space) in the health queries.

Usage of Query Operators and Special Characters

In search queries, query operators ("and", "or", "not", etc) are used to formulate complex queries. In this study, we considered the following operators: AND, OR, +, &, other (NOT, AND NOT, OR NOT, & NOT). Special characters are characters apart from letters (a-z) and digits (0-9). The significance of special characters in a health search query depends on the usage

of special characters in the medical domain. For example, OHIS may mention values in different formats, eg, 2.3 ml, 40%, 17-19, or $200 (for the cost of a drug or procedure). We analyzed the usage of search query operators and special characters in health queries based on their usage frequency in the PCs and SDs search queries.

Misspellings in Health Queries

OHISs occasionally make spelling mistakes while searching for health information. To analyze the frequency of such errors, we used a dictionary-based approach. We first generated a dictionary of words using the Zyzzyva wordlist [49], the Hunspell dictionary [50], and its medical version (OpenMedSpell [51]), comprising a total of 275,270 unique words. We used this dictionary to check misspellings in health search queries from PCs and SDs.

Type of Search Queries

OHISs express their health information need by formulating health search queries on Web search engines. In general, each health search query indicates some health information need. OHISs can express their information need either by formulating search queries using keywords or asking questions (wh-questions and yes/no questions). For this analysis, we considered the following wh-questions (lexicon): "What", "How", "?", "When", "Why", and others ("Who" "Where", "Which"). Note that although "?" does not come under the wh-questions category, we have included it for simplicity. Yes/No questions are usually used to check factual information; for example, whether coffee

is bad for the heart. In this analysis, we considered yes/no questions that start with "Can", "Is", "Does", "Do", "Are", and others ("Could", "Should", "Will", "Would"). Using the lexicon for wh-questions and yes/no questions, we performed text analysis on the search queries from PCs and SDs to count the number of queries with wh-questions and yes/no questions. Search queries that do not contain any question (wh- or yes/no) are classified as keyword-based. Additionally, for different wh- and yes/no questions, we computed their usage frequency in search queries from PCs and SDs.

Linguistic Analysis of Health Queries

Previous studies in generic search query log analysis have identified that understanding the linguistic structure, including phrase identification, entity spotting and discriptiveness (level of context), of search queries can improve Web Information Retrieval systems [42,43]. However, these efforts have not been applied extensively to health search queries, and hence in order to understand the linguistic structure of health queries, we performed part-of-speech analysis on search queries using Stanford's POS tagger [52]. For this analysis, we considered nouns, verbs, adjectives and adverbs. We mapped all subtypes in part-of-speech (eg, proper nouns, common nouns, compound nouns) to the main part-of-speech (eg, nouns). We analyzed the usage of different part-of-speech types in health queries based on their usage frequency in the PCs and SDs search queries.

Results

Top Health Queries

Most of the top search queries from both PCs and SDs are for symptom descriptions (eg, "lupus symptoms"). Another common way an OHIS searches for health information is by disease name (eg, "Lupus"). Chronic diseases (cancer, cardiovascular disease, diabetes) and diet (Mediterranean diet, gluten free food) are also searched often. Based on the top 100 search queries from PCs and SDs, we found that 48.49% of the search queries are different between PCs and SDs. However, due to the Mayo Clinic business confidentiality, we are not in a position to disclose the actual top search queries and numbers publicly.

Health Categories

While searching for health information, one in every three OHIS searches for "Symptoms" (Figure 2). Other popular health categories are "Causes" and "Treatments & Drugs". Our analysis shows that the distribution of search queries for different health categories differs with the device used for the health search. At the same time, both PCs and SDs follow a similar pattern for distribution of the search queries between health categories. The percentage of OHIS searching for "Symptoms" is higher from SDs as compared to that from PCs. While for other health categories, the percentage of queries from PCs is slightly higher than that of SDs. Interestingly, one of the least searched health categories is "Prevention".

Figure 2. Distribution of the search queries by health categories.
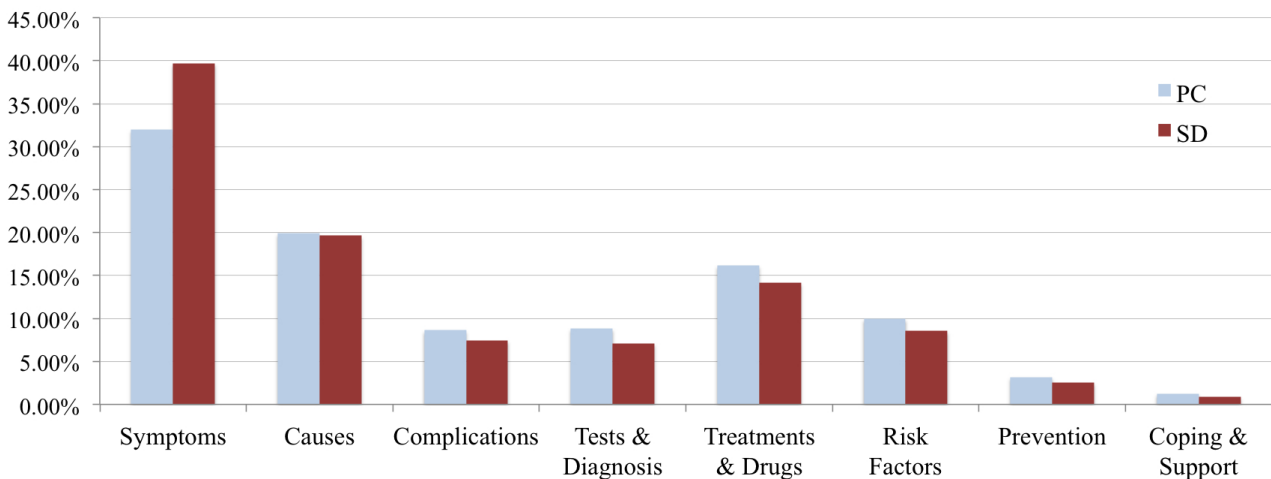
Categorization Based on the Information Mentioned in the Health Queries

The following are some of our observations based on the information referenced in the search queries (Table 1). The data indicate that the number of search queries mentioning words related to women's health is considerably higher compared to that of men. This implies that OHIS search for health information specifying women more often. The percentage of OHIS who use words related to "woman" in search queries is higher for PCs compared to SDs. Considering age group–related

search queries, more than 60% of the queries are related to children. The percentage of OHIS that mention terms related to children in search queries is much higher for SDs compared to PCs. When considering a mention of the time of day in search queries, terms related to "Night" are mentioned most often (>60%) followed by words related to "Morning". Very few search queries have words related to "Afternoon" and "Evening". The percentage of OHIS using words related to "Morning" in search queries is higher for SDs compared to PCs, while the percentage of OHIS mentioning words related to "Night" in search queries is higher for PCs.

Table 1. Categorization of health search queries based on the information mentioned in the queries such as gender, age group, and temporal information (June 2011-May 2013).

		Personal computers		Smart device	
		QwoR %	QWR %	QwoR %	QWR %
Gender					
	Men	25.62	17.25	28.48	18.17
	Women	74.38	82.75	71.52	81.83
Age group					
	Children	66.55	59.60	79.33	74.39
	Teen	7.25	5.08	3.69	2.37
	Adults	18.60	31.68	13.64	21.72
	Elders	7.60	3.64	3.34	1.52
Temporal					
	Morning	26.85	29.93	31.93	39.14
	Afternoon/Evening	5.84	4.39	4.10	1.73
	Night	67.31	65.68	63.98	59.13

Health Query Length

The average search query length (Figures 3 and Figure 4) for QwoR (PCs: 4.82 words and 26.73 characters; SDs: 5.33 words and 27.41 characters) is much larger than the average length of QwR (PCs: 2.90 words and 17.61 characters; SDs: 3.29 words and 18.86 characters). This indicates that longer search queries result in fewer repetitions, while shorter queries tend to be repeated more often. The analysis, although derived from a limited dataset, implies that in general health search queries tend to be longer than general search queries (not specific to one domain), as the average length of general search query from PCs is 2-2.35 words [53-55] and from SDs is 2.3 words [56]. This potentially indicates that OHISs describe their health information needs in more detail by adding relevant health context to the search query. Surprisingly, the average length of search query from SDs for both QwoR and QwR is slightly larger than queries from PCs.

Figure 3. Distribution of the search queries by number of words and number.
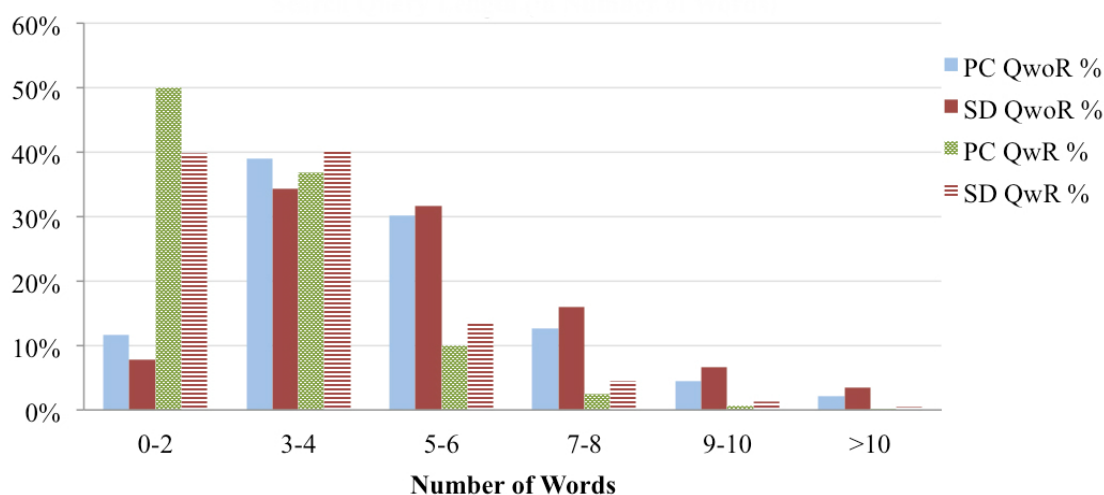
Figure 4. Distribution of the search queries by number of characters.
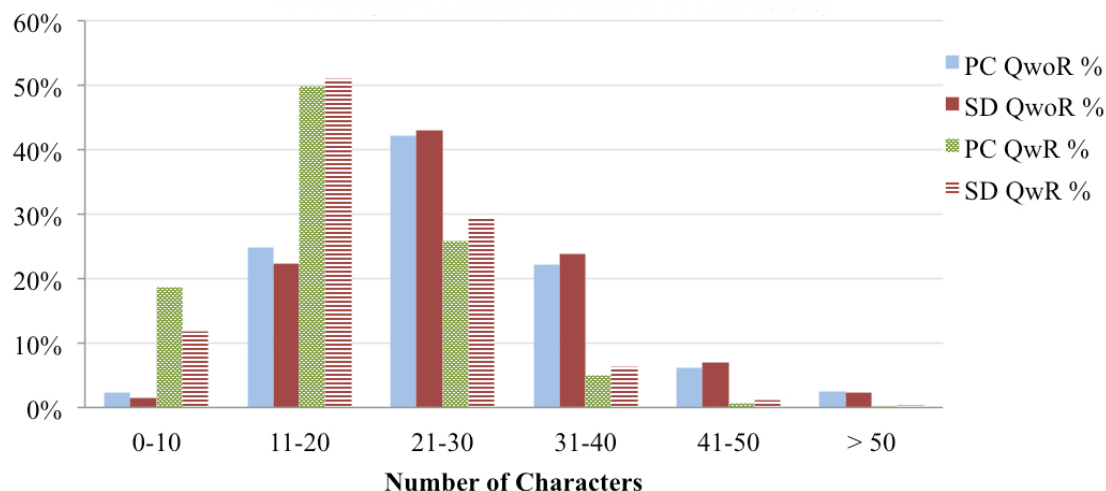
Usage of Query Operators and Special Characters

In considering both PCs and SDs, approximately 10% of QwoR and 3% of QwR use at least one query operator. For QwR, the percentage of OHIS who use query operators in search queries is higher for SDs than PCs, while in the case of QWOR it is higher for PCs. AND is the most popular operator, followed by OR and "+". Overall variations of "and" (AND, &, +) operators comprise more than 90% of operator usage. Considering QwoR, OHIS use AND OR query operators more often from SDs than that from PCs. Considering both PCs and SDs, around 10% of QwoR and 4% of QwR have at least one special character (Table 2). The percentage of OHIS using special characters in search queries is higher for PCs compared to SDs.

Table 2. Usage of query operators and special characters (June 11-May 13).

		Personal computers		Smart device	
		QwoR %	QwR %	QwoR %	QwR %
Number of operators					
	0	90.08	97.35	90.23	96.53
	>0	9.92	2.65	9.77	3.47
Query operators usage					
	AND	78.96	86.53	82.01	85.05
	+	11.24	4.37	6.29	3.08
	OR	6.95	5.20	8.74	6.78
	&	2.63	1.42	2.57	1.28
	Other	0.24	2.49	0.40	3.82
Special characters					
	0	89.02	95.66	90.54	96.72
	>0	10.98	4.34	9.46	3.29
Spelling mistake					
	0	68.21	87.47	69.07	87.88
	>0	31.80	12.54	30.94	12.12

Misspellings in Health Queries

For QwoR and QwR, approximately 31% and 12% of queries, respectively, have at least one spelling mistake (Table 2). OHISs make slightly more spelling mistakes while searching health information from PCs than SDs.

Types of Health Queries

As indicated by the analysis in Figure 5, OHISs predominantly formulate search queries using keywords, though wh-questions and yes/no questions are also substantial. Considering QwoR, OHISs ask more (wh- and yes/no) questions from SDs than PCs. In wh-questions (Figure 6), OHISs mostly use "What" and "How" in the search queries, and both of them generally signify that more descriptive information is needed. OHISs ask more temporal questions ("When") using SDs than PCs, while OHISs

ask more "What" questions using PCs than SDs. In yes/no questions (Figure 7), OHISs generally start search queries with "Can," "Is", and "Does". OHISs ask more yes/no questions starting with "Can" using SDs than using PCs, while the percentage of questions starting with "Is" and "Does" comes more from PCs.

Figure 5. Types of health search queries (how health information need is expressed).
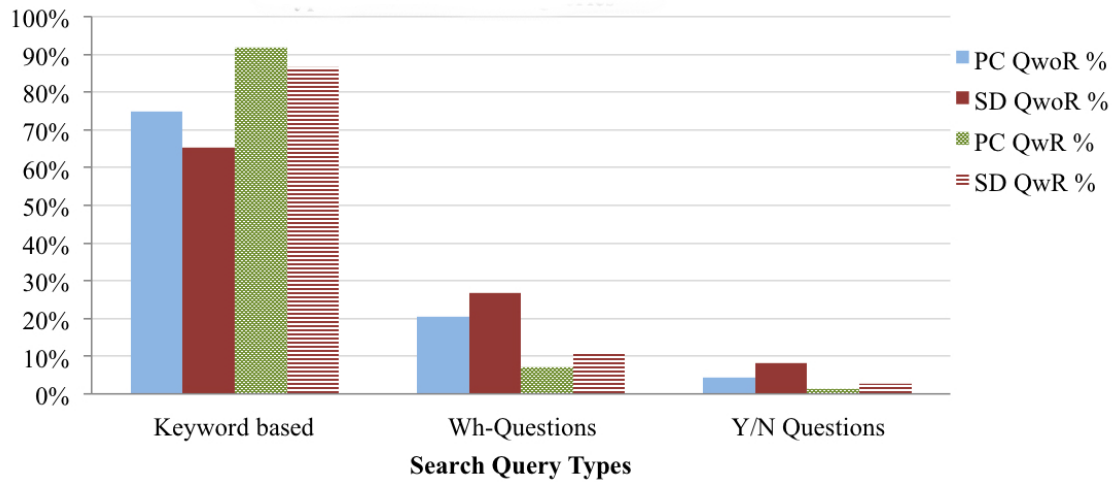
Figure 6. Distribution of the search queries based on type of wh-questions.
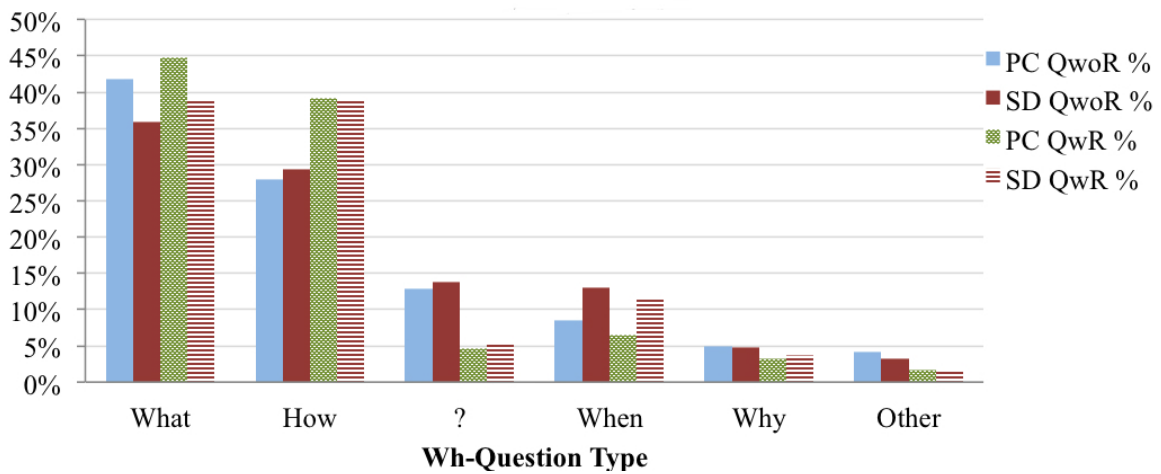
Figure 7. Distribution of the search queries based on type of yes/no questions.
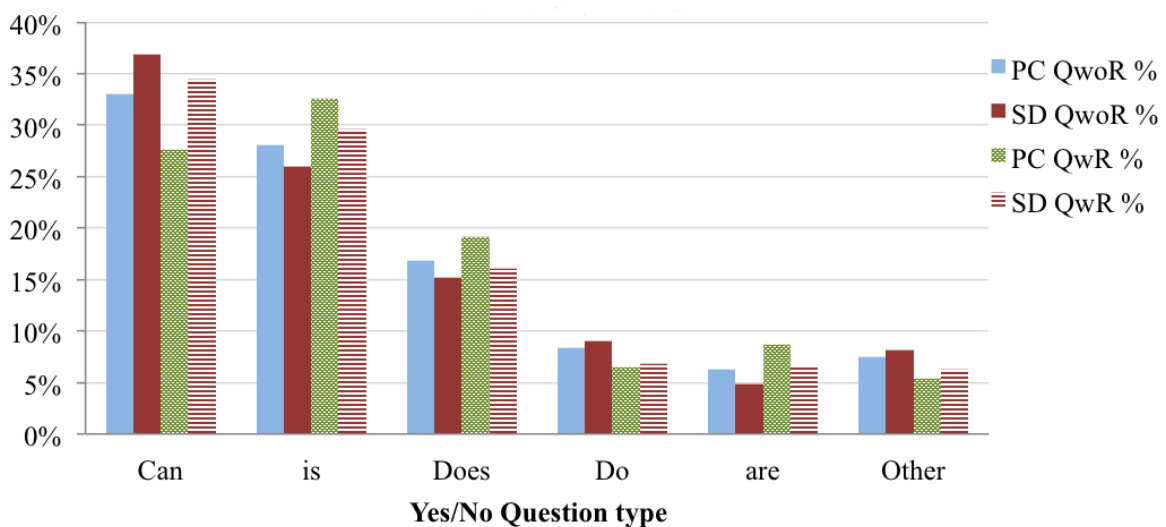
Linguistic Analysis of Health Queries

In health search queries, nouns typically denote entities like disease names, health categories, etc. Almost all health search queries have at least one noun. In the case of QwR, most of the search queries (>70%) have 1-2 nouns, while in the case of QwoR, most of the search queries (>60%) have 2-3 nouns. There is no considerable difference in noun usage between PCs and SDs. A verb conveys an action or an occurrence, for example "how to control (*verb*) diabetes (*noun*)?". Considering QwoR, OHIS use at least one verb in 37% of queries from PCs and 47% in queries from SDs. Adverbs are words that modify a verb, an adjective, and another adverb, while an adjective is a "describing" word, giving more information about the object signified; for example, "extremely (*adverb*) bad (*adjective*) stomach (*noun*) pain (*noun*)". Very few search queries have at least one adverb. Considering QwoR, 45.66% of the queries from PCs and 48.50% of the queries from SDs have at least one adjective. This indicates that the percentage of search queries with at least one verb/adverb/adjective is higher for SDs than for PCs (see Table 3).

Table 3. Linguistic analysis of health search queries (June 2011-May 2013).

		Personal computers		Smart device	
		QwoR %	QwR %	QwoR %	QwR %
Nouns					
	0	0.96	3.19	1.11	1.67
	1	14.31	28.17	14.52	26.93
	2	36.01	46.87	36.97	47.38
	3	31.34	17.75	31.61	19.79
	>3	17.37	4.01	15.80	4.23
Verb					
	0	62.96	83.34	53.09	78.96
	>0	37.04	16.66	46.92	21.05
Adverb					
	0	93.86	95.56	91.01	95.38
	>0	6.15	4.45	9.00	4.62
Adjective					
	0	54.32	69.71	51.51	66.14
	>0	45.68	30.30	48.50	33.87

Discussion

Overview

Increasingly, individuals are actively participating in learning and managing their health by leveraging online resources. The percentage of people using the Internet and the usage of smart devices for health information searching is increasing rapidly. PCs and SDs have very distinct characteristics in terms of readability, user experience, accessibility, etc. These distinct characteristics provide some pros and cons for PCs and SDs: Web browsing and readability are better on PCs while accessibility is better for SDs. Also socioeconomic factors, such as age, gender, income level, education, familiarity with new technologies and devices [4,9], play an important role in the usage of PCs and SDs in general and for online health information seeking. Device characteristics and socioeconomic differences in device usage have an effect on OHISB [4,5,9]. Therefore, in order to improve the health information searching process, it is necessary to understand both aspects, that is, how an OHIS searches for health information and how device choice influences online health information seeking.

In this study, we performed a comparative analysis on the most frequent health search queries launched from PCs and SDs to understand the effects of device type (PCs vs SDs) used for online health information seeking. The analysis dataset consists of search queries between June 2011 and May 2013, which were submitted from Web search engines and directed OHISs to the Mayo Clinic website. The website is visited by millions of unique OHIS every day, and it offers an identical appearance and accessibility for both PCs and SDs using standard Web search engines and Web browsers.

Principal Results

Following are some of the insights that surfaced from this study. Most of the top search queries from both PCs and SDs are related to symptoms, health conditions, chronic diseases, and diet. Our top search query analysis indicates that the device used has a significant effect on health information searching and the health information searched via different devices is also different (48.49%). While searching for health information, one in every three OHISs searches for "Symptoms". Other popular health categories that OHISs search for are "Causes" and "Treatments & Drugs". The analysis suggests that the distribution of search queries for different health categories

differs with the device used for health search. Even though most of the diseases can be prevented with some lifestyle and diet changes, very few OHIS search for *preventive* health information. This highlights the fact that we need to promote preventive health care more vigorously.

While searching for health information, OHISs specify words related to women and children more often than that of men and any other age group. The higher percentage of women seeking online health information could be a reason [4,5]. The percentage of OHISs who use words related to "women" and "night" in search queries is higher for PCs than for SDs, while "children" and "morning" are higher for SDs compared to PCs. Health search queries are longer than general search queries, which implies that OHISs describe health information need in more detail. Longer search queries also denote OHIS's interest in more specific information about the disease; subsequently, OHISs use more words to narrow down to a particular health topic. The average health search query length from SDs is longer than that of PCs, and while typing on SDs is slower and more difficult than typing on PCs, we posit that OHISs might be relying more on Web search engines' auto-completion functionality, as well as on most devices' speech recognition facilities, which might be increasing the length of search queries from SDs as compared to that from PCs. These results highlight the differences between usage of PCs and SDs for online health information seeking. The findings can be used by health websites and health application developers to better understand OHISB for PCs and SDs, understand OHIS's health information needs, and better organize health information content for PCs and SDs users.

For PCs and SDs, 1 in 3 QwoR, and 1 in 10 QwR contained at least one spelling mistake. These mistakes place a burden on the search process and may lead users to incorrect or irrelevant information. The search engine's auto-completion feature, spelling correction/suggestion, and devices' speech recognition facilities might be contributing to reducing misspelled words in search queries. Almost all health search queries have at least one noun. In addition to nouns, OHISs use verbs, adverbs, and adjectives while formulating search queries to provide more context for the topic of interest. The percentage of search queries with at least one verb/adverb/adjective is higher for SDs as compared to PCs. This implies that health search queries from SDs are more descriptive as compared to queries from PCs. OHISs formulate search queries by using keywords most frequently, followed by wh-questions and yes/no questions. Considering QwoR, OHIS ask more questions via SD than PC. In wh-questions, OHISs mostly use "What" and "How" in search queries, and both of them generally signify a need for more descriptive information while search queries in the form of yes/no questions indicate interest in factual information.

Since search queries are a fundamental part of health information searching, it is essential that we understand characteristics of health search queries and the role of the device used for searching. This study provides useful insights for online health information retrieval systems. The linguistic structure of a search query has implications in information retrieval using Web search engines [42,43]. Cory Barr et al [42] highlight the importance of recognizing part-of-speech information of the input search

query to improve search results and demonstrate that the part-of-speech is a significant feature for information retrieval. Our study provides distribution of part-of-speech in health search queries from PCs and SDs. Expressiveness or descriptiveness of the search queries has a significant impact on quality of the search results using Web search engines [43]. Phan et al [57] specify that with the increase in search query length, the descriptiveness of the query increases. Our study gives basic understanding about health search query descriptiveness based on health query length and part-of-speech analysis. Previous research in information retrieval have identified various important features of search queries such as usage of search query operators [58], misspellings, query length [53-57], query type (keyword-based, wh-questions, yes/no questions), and part-of-speech [42,43]. We presented a comprehensive analysis of these features for health search queries via PCs and SDs.

Comparison With Related Work

This study contributes a comparative analysis performed on large-scale health search queries to understand the effects of device type (SDs vs PCs) used on OHISB. As discussed in the "Background and Significance" section, previous efforts have used several approaches to understand OHISB including (1) focus groups and user surveys, and (2) analyzing health-related Web search query logs. To the best of our knowledge, there is not much research on understanding the effect of devices on online health search behavior. In our work, we bridge this knowledge gap by analyzing more than 100 million health search queries from PCs and SDs to understand how device choice influences online health information seeking. In addition, we presented analysis for both QwR and QwoR in order to avoid bias from queries with a high number of repetitions. Moreover, we analyzed linguistic structure of health search queries from PCs and SDs, which has implications for Web search engines and information retrieval systems [42,43].

Limitations

The results of this study are derived from analysis limited to health search queries from Web search engines that led users to Mayo Clinic website. Even though Mayo Clinic web pages often ranked high in Web search engines, not all health information seekers visited the Mayo Clinic website. Also, this analysis is based on the top one million health queries per month (PCs/SDs) rather than the entire health traffic to Mayo Clinic site. In this work, we considered search queries from smartphones and tablets into same categories (ie, smart devices) as the search queries are differentiated based on the operating system of the device used for search, and not the type of specific device per se (eg, Apple iPhone vs iPad vs Android phone). The focus of this study is limited to analysis of a search query log, and we have not analyzed associated socioeconomic factors due to anonymized nature of the data. Previous studies have identified that socioeconomic factors such as age, gender, education, and income have an effect on device usage and OHISB [4,5,9]. Further research in analyzing health search queries based on socioeconomic factors can extend our knowledge about how socioeconomic factors affect health search query formation and the type of health information searched.

XSL•FO

RenderX

Future Work

In the future, we will extend this work by performing a semantic analysis on the data using biomedical knowledge bases and ontologies. Specifically, we plan to leverage insights from this work and use semantic Web technologies to facilitate health search experience by developing more advanced next-generation knowledge and content delivery systems. Semantic analysis in combination with advanced natural language processing techniques will help us acquire a deeper understanding of OHISB.

Conclusions

We presented a comprehensive analysis of large-scale health search queries from personal computers (desktops/laptops) and smart devices (smartphones/tablets) in order to understand the effects of device type on online health information search behavior. We noted that online *health* information search behavior differs from *general* online information search. Also, the type of device used for online health information search plays an important role and alters the health information search behavior. A greater understanding of OHIS's needs, especially how they search and what they search for, may help us understand behavioral changes that will lead to improvement in online health information seeking and a more balanced approach to wellness and prevention. This study extends our knowledge about online health information search behavior and provides useful information for Web search engines, health-centric websites, health care providers, and health care–centric application developers. Finally, we anticipate that this work will help empower OHISs in their quest for health information and facilitate their health information search efforts by enabling the development of more advanced next-generation knowledge and content delivery systems.

Conflicts of Interest

None declared.

References

1. Berkman ND, Davis TC, McCormack L. Health literacy: what is it? J Health Commun 2010;15 Suppl 2:9-19. [doi: 10.1080/10810730.2010.499985] [Medline: 20845189]

2. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]

3. Tustin N. The role of patient satisfaction in online health information seeking. J Health Commun 2010 Jan;15(1):3-17. [doi: 10.1080/10810730903465491] [Medline: 20390974]

4. Fox S, Duggan M. Pew Internet & American Life Project. 2013. Health online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2014-06-25] [WebCite Cache ID 6QbBAkxys]

5. Higgins OSJ, Barry MM, Domegan C. A literature review on health information seeking behaviour on the web: a health consumer and health professional perspective. ECDC 2011:1-11.

6. Della Mea V. What is e-health (2): the death of telemedicine? J Med Internet Res 2001;3(2):E22 [FREE Full text] [doi: 10.2196/jmir.3.2.e22] [Medline: 11720964]

7. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]

8. Duggan M, Smith A. Pew internet & American Life Project. 2013. Cell Internet use 2013 URL: http://www.pewinternet.org/2013/09/16/cell-internet-use-2013/ [accessed 2014-06-25] [WebCite Cache ID 6QbBLNZJU]

9. Fox S, Duggan M. Pew internet & American Life Project. 2012. Mobile health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [accessed 2014-06-25] [WebCite Cache ID 6QbBQPBEk]

10. Roto V. Web browsing on mobile phones - characteristics of user experience. Espoo, Finland: Helsinki University of Technology; 2006.

11. IBM Web Analytics Tool. URL: http://www-03.ibm.com/software/products/en/on-premise-web-analytics/ [accessed 2014-02-02] [WebCite Cache ID 6NXP32XzE]

12. Mayo Clinic's consumer health information website. URL: http://www.mayoclinic.org/ [accessed 2014-02-02] [WebCite Cache ID 6NXP934Iu]

13. Lorence DP, Park H, Fox S. Assessing health consumerism on the Web: a demographic profile of information-seeking behaviors. J Med Syst 2006 Aug;30(4):251-258. [Medline: 16978004]

14. Drentea P, Goldner M, Cotten S, Hale T. The association among gender, computer use and online health searching, and mental health. Information, Communication & Society 2008;11(4):509-525.

15. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]

16. Weaver JB, Mays D, Weaver SS, Hopkins GL, Eroglu D, Bernhardt JM. Health information-seeking behaviors, health indicators, and health risks. Am J Public Health 2010 Aug;100(8):1520-1525. [doi: 10.2105/AJPH.2009.180521] [Medline: 20558794]

17. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]

18. Wangberg SC, Andreassen HK, Prokosch HU, Santana SM, Sørensen T, Chronaki CE. Relations between Internet use, socio-economic status (SES), social support and subjective health. Health Promot Int 2008 Mar;23(1):70-77 [FREE Full text] [doi: 10.1093/heapro/dam039] [Medline: 18083686]

19. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]

20. Flynn KE, Smith MA, Freese J. When do older adults turn to the internet for health information? Findings from the Wisconsin Longitudinal Study. J Gen Intern Med 2006 Dec;21(12):1295-1301 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00622.x] [Medline: 16995892]

21. White RW, Horvitz E. From health search to healthcare: explorations of intention and utilization via query logs and user surveys. J Am Med Inform Assoc 2014;21(1):49-55. [doi: 10.1136/amiajnl-2012-001473] [Medline: 23666794]

22. White RW, Horvitz E. Cyberchondria: studies of the escalation of medical concerns in web search. ACM Trans. Inf. Syst 2009 Nov 01;27(4):1-37. [doi: 10.1145/1629096.1629101]

23. White RW, Horvitz E. Experiences with web search on medical concerns and self diagnosis. AMIA Annu Symp Proc 2009;2009:696-700 [FREE Full text] [Medline: 20351943]

24. White R, Horvitz E. From web search to healthcare utilization: privacy-sensitive studies from mobile data. J Am Med Inform Assoc 2013 Jan 1;20(1):61-68 [FREE Full text] [doi: 10.1136/amiajnl-2011-000765] [Medline: 22661560]

25. Zhang Y, Wang P, Heaton A, Winkler H. Health information searching behavior in MedlinePlusthe impact of tasks. 2012 Presented at: 2nd ACM SIGHIT International Health Informatics Symposium, ACM; January 28-30, 2012; Miami, FL.

26. Eysenbach G, Köhler C. Health-related searches on the Internet. JAMA 2004 Jun 23;291(24):2946. [doi: 10.1001/jama.291.24.2946] [Medline: 15213205]

27. Cartright MA, White RW, Horvitz E. Intentions and attention in exploratory health search. 2011 Presented at: 34th international ACM SIGIR conference on Research and development in Information Retrieval; July 24-28, 2011; Beijing, China p. 65-74. [doi: 10.1145/2009916.2009929]

28. White RW, Drucker SM. Investigating behavioral variability in web search. 2007 Presented at: 16th international conference on World Wide Web, ACM; May 8-12, 2007; Banff, Alberta, Canada p. 21-30. [doi: 10.1145/1242572.1242576]

29. Zeng QT, Kogan S, Plovnick RM, Crowell J, Lacroix EM, Greenes RA. Positive attitudes and failed queries: an exploration of the conundrums of consumer health information retrieval. Int J Med Inform 2004 Feb;73(1):45-55. [doi: 10.1016/j.ijmedinf.2003.12.015] [Medline: 15036078]

30. Spink A, Yang Y, Jansen J, Nykanen P, Lorence DP, Ozmutlu S, et al. A study of medical and health queries to web search engines. Health Info Libr J 2004 Mar;21(1):44-51. [doi: 10.1111/j.1471-1842.2004.00481.x] [Medline: 15023208]

31. Yang CC, Winston F, Zarro MA, Kassam-Adams N. A study of user queries leading to a health information website: AfterTheInjury. org. : ACM; 2011 Presented at: iConference; 2011; Seattle, Washington p. 267-272.

32. Eysenbach G, Kohler C. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. AMIA Annu Symp Proc 2003:225-229 [FREE Full text] [Medline: 14728167]

33. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]

34. Ocampo AJ, Chunara R, Brownstein JS. Using search queries for malaria surveillance, Thailand. Malar J 2013;12:390 [FREE Full text] [doi: 10.1186/1475-2875-12-390] [Medline: 24188069]

35. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. N Engl J Med 2009 May 21;360(21):2153-5, 2157 [FREE Full text] [doi: 10.1056/NEJMp0900702] [Medline: 19423867]

36. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564 [FREE Full text] [doi: 10.1086/630200] [Medline: 19845471]

37. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248 [FREE Full text] [Medline: 17238340]

38. 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;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]

39. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. J Med Internet Res 2013;15(7):e147 [FREE Full text] [doi: 10.2196/jmir.2740] [Medline: 23896182]

40. Herskovic JR, Tanaka LY, Hersh W, Bernstam EV. A day in the life of PubMed: analysis of a typical day's query log. J Am Med Inform Assoc 2007;14(2):212-220 [FREE Full text] [doi: 10.1197/jamia.M2191] [Medline: 17213501]

41. Islamaj Dogan R, Murray GC, Névéol A, Lu Z. Understanding PubMed user search behavior through log analysis. Database (Oxford) 2009;2009:bap018 [FREE Full text] [doi: 10.1093/database/bap018] [Medline: 20157491]

42. Barr C, Jones R, Regelson M. The linguistic structure of English web-search queries. 2008 Presented at: Conference on Empirical Methods in Natural Language Processing; 25-28 October, 2008; Waikiki, Honolulu, Hawaii p. 1021-1030.

43. Croft WB, Metzler D, Strohman T. Search engines: Information retrieval in practice. New York: Addison-Wesley Reading; 2010.

44. MedlinePlus - Health Information from the National Library of Medicine. URL: http://www.nlm.nih.gov/medlineplus/ [accessed 2014-02-02] [WebCite Cache ID 6NXPIiNBS]

45. WebMD - Better information Better health. URL: http://www.webmd.com/ [accessed 2014-02-02] [WebCite Cache ID 6NXPNb47D]

46. Dictionary.com. URL: http://dictionary.reference.com/ [accessed 2014-02-02] [WebCite Cache ID 6NXPU2Ma1]

47. Merriam Webster. URL: http://www.merriam-webster.com/ [accessed 2014-02-02] [WebCite Cache ID 6NXTfhav8]

48. Oxford Dictionaries. URL: http://www.oxforddictionaries.com/us [accessed 2014-02-02] [WebCite Cache ID 6NXPdvmwD]

49. Zyzzyva: The Last Word in Word Study. URL: http://www.zyzzyva.net/wordlists.shtml [accessed 2014-02-02] [WebCite Cache ID 6NXPhfsmi]

50. Hunspell dictonary. URL: http://hunspell.sourceforge.net/ [accessed 2014-02-02] [WebCite Cache ID 6NXPoLaXA]

51. OpenMedSpel for Hunspell. URL: http://www.e-medtools.com/Hunspel_openmedspel.html [accessed 2014-02-02] [WebCite Cache ID 6NXPsL29d]

52. Toutanova K, Klein D, Manning CD, Singer Y. Feature-rich part-of-speech tagging with a cyclic dependency network. 2003 Presented at: NAACL '03 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology; May 27 - June 1, 2003; Edmonton, Canada p. 173-180. [doi: 10.3115/1073445.1073478]

53. Silverstein C, Marais H, Henzinger M, Moricz M. Analysis of a very large web search engine query log. ACM SIGIR Forum 1999;33(1):6-12.

54. Spink A, Wolfram D, Jansen MB, Saracevic T. Searching the web: The public and their queries. Journal of the American Society for Information Science and Technology 2001;52(3):226-234.

55. Wikipedia article: Web search query. URL: http://en.wikipedia.org/wiki/Web_search_query [accessed 2014-02-02] [WebCite Cache ID 6NXTaM9lr]

56. Kamvar M, Baluja S. A large scale study of wireless search behavior: Google mobile search. 2006 Presented at: CHI '06 Conference on Human Factors in Computing Systems; April 22-27, 2006; Montreal, Quebec, Canada p. 701-709.

57. Phan N, Bailey P, Wilkinson R. Understanding the relationship of information need specificity to search query length. 2007 Presented at: 30th annual international ACM SIGIR conference on Research and development in information retrieval; 23-27 July, 2007; Amsterdam p. 709-710.

58. Eastman CM, Bernard JJ. Coverage, relevance, and ranking: The impact of query operators on Web search engine results. ACM Transactions on Information Systems (TOIS) 2003;21(4):383-411.

Abbreviations

OHIS: online health information seeker
OHISB: online health information seeking behavior
PC: personal computer (desktop, laptop)
QwoR: queries without considering repetition counts
QwR: queries with considering repetition counts
SD: smart device (smartphone, tablet)

XSL•FO
RenderX

Letter to the Editor

Consensus on Use of the Term "App" Versus "Application" for Reporting of mHealth Research

Thomas Lorchan Lewis[1*], BSc (Hons); Matthew Alexander Boissaud-Cooke[1*], BMedSc; Timothy Dy Aungst[2], PharmD; Gunther Eysenbach[3,4], MD, MPH, FACMI

[1]Warwick Medical School, University of Warwick, Coventry, United Kingdom

[2]Department of Pharmacy, MCPHS University, Worcester, MA, United States

[3]JMIR Publications, Toronto, ON, Canada

[4]University Health Network, Techna Institute, and University of Toronto, Toronto, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Thomas Lorchan Lewis, BSc (Hons)
Warwick Medical School
University of Warwick
Medical Teaching Centre
Gibbet Hill Road
Coventry, CV4 7AL
United Kingdom
Phone: 44 7876453511
Fax: 44 2476528375
Email: t.lewis@warwick.ac.uk

Letter

We read with great interest a number of recent articles from the Journal of Medical Internet Research (JMIR) published by Bierbrier et al [1], Kim et al [2], and Choi et al [3], investigating different aspects of mHealth technology, specifically medical "applications" or "apps". Many individuals, researchers, academic institutions, and other professional bodies often use these terms interchangeably. It is apparent that there is no clear consensus for which term should be used as both are recognized terms for a software program designed to run on smartphones, tablet computers and other mobile devices.

As the field continues to expand, we believe that the inconsistent use of terminology used may present a problem for future researchers to systematically identify and conduct appropriate literature searches. Figure 1 shows a graph of the cumulative number of PubMed search results by year related to keywords relating to "medical application" and "app" respectively since 1975 (Figure 1). This shows the clear exponential growth in this field as the amount of research in this field continues to grow. In particular it is worth drawing attention to the fact that the term "medical application" is used considerably more often in a number of medical specialties indicating it is not specific to mobile heath. The inconsistent use of terminology is also apparent in the use of keywords and Medical Subject Headings

(MeSH) terms. It is currently unclear which are the most appropriate keywords for selection with many researchers using a variety of terms, common examples include: mobile device vs smartphone vs cellphone, mobile tablet vs mobile computer, and applications vs apps.

We believe it is now time for the mHealth research community to come to a universal consensus on whether studies should refer to medical "apps" or "applications". Standardization of terminology will enable researchers and other health care professionals to:

- identify relevant articles and improve the literature search process through use of common search terms across different modalities;
- identify common MeSH terms which describe interventions utilizing mobile medical software (which is currently lacking);
- ensure databases categorize mobile health interventions more effectively for future researchers;
- improve the distinction between software designed for use on mobile devices and desktop devices;
- improve reporting of studies investigating mobile health interventions.

We believe that we should use the term "app" [plural "apps"], rather than "applications" for the following reasons:

XSL•FO
RenderX

1. this is the commonly used term in the public domain, social media and by the major hardware manufacturers;
2. the term medical application is not specific enough to software mHealth programs utilized on mobile devices. Many other medical research fields utilize the term "medical applications" whilst the term "app" is more specific to software programs utilized on mobile devices;
3. there is currently no MeSH keyword for the term "mobile medical app" while there are already preexisting MESH keywords related to "medical informatics application" or "mobile app", neither of which is specific for mHealth interventions;
4. use of the term "application" may be misleading particularly for lay users who may believe this term represents software designed for desktop computers.

In conclusion, we would recommend that leading eHealth medical informatics publications such as the JMIR journals implement a policy to utilize common nomenclature moving forward to facilitate improved reporting of studies investigating mobile medical app interventions.

Thomas Lorchan Lewis, Matthew Alexander Boissaud-Cooke, Timothy Dy Aungst

Figure 1. A graph showing cumulative number of PubMed search results by year since 1975 for keywords related to mobile applications (search carried out on April 7, 2014).
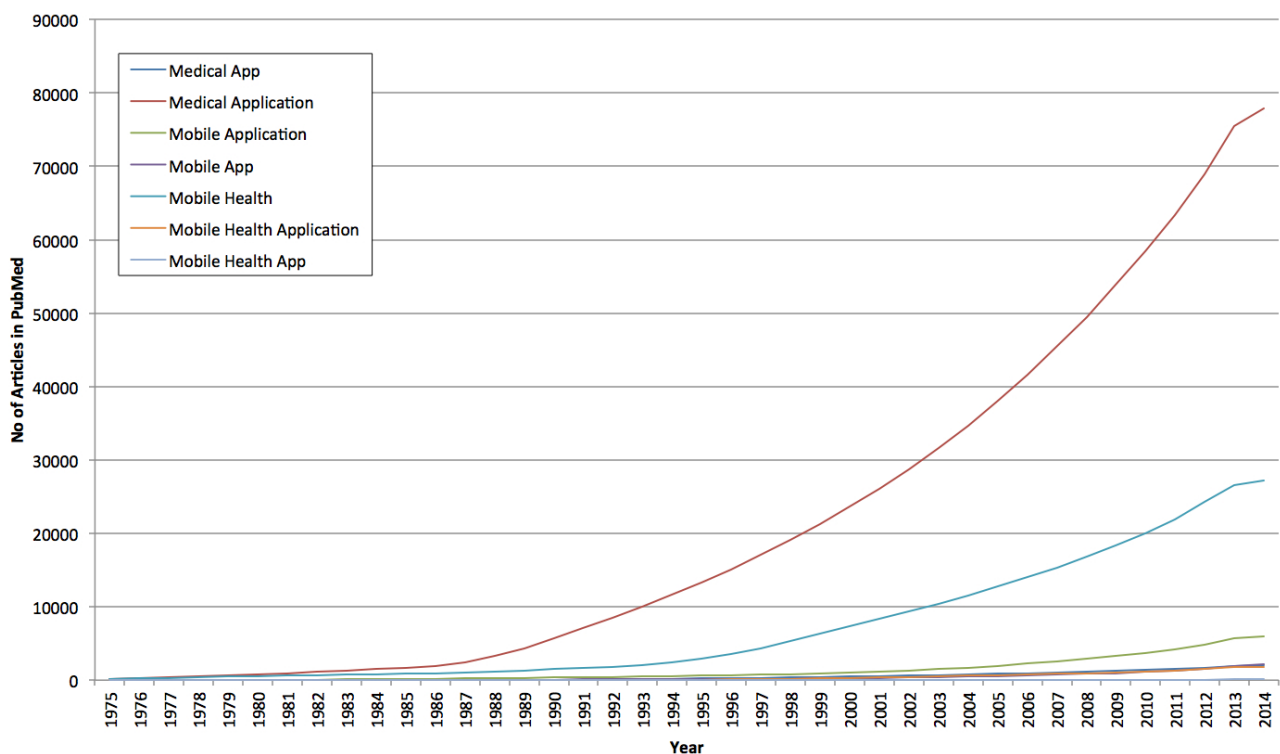
Editorial Response

We appreciate the thoughtful and important comment of Lewis et al and fully agree about the need of a consistent terminology for mobile apps, as well as the preferential use of the term "app". In fact, JMIR Publications maintains an internal style guide, and already in June 2013 introduced a new guideline for our copyeditors where we explicitly ask to enforce use of the word "app" rather than "application", even though the word "app" was originally a short form of "application software". The use of the word "application" in the title of the paper cited by Lewis et al [2] was an oversight on the part of the freelance copyeditor assigned to the manuscript, and we will be more vigilant in the future to enforce the term "app". We also have other standards which should help indexing and retrieval in particular in the context of systematic reviews. For example, we prefer the term "mobile phone" over "smartphone" in title and abstract, as the latter is often forgotten by systematic reviewers searching for "mobile" technology studies (we also noted it is not mentioned in Figure 1). In addition, all papers referring to mobile technologies are indexed with the theme keyword "mhealth". These policies extend to all JMIR journals, including Journal of Medical Internet Research, JMIR Research Protocols, JMIR mHealth and uHealth, JMIR Serious Games, JMIR Medical Informatics, JMIR Human Factors, JMIR Mental Health, interactive Journal of Medical Research, Medicine 2.0 and others. We hope that other journals will follow and adopt these terminology standards, which should ultimately also make it into reporting guidelines such as CONSORT EHEALTH [4] .

G. Eysenbach, Editorial Director, JMIR Publications

Conflicts of Interest

TLL and TDA are writers and editors for iMedicalApps.com, a website dedicated towards providing news on the integration of mobile technology into medical care and the reviewing of medical apps for mobile devices. They do not consult nor receive reimbursement from app developers or creators.

References

1. Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. J Med Internet Res 2014;16(2):e32 [FREE Full text] [doi: 10.2196/jmir.3062] [Medline: 24491911]
2. Kim JH, Kwon SS, Shim SR, Sun HY, Ko YM, Chun DI, et al. Validation and reliability of a smartphone application for the international prostate symptom score questionnaire: a randomized repeated measures crossover study. J Med Internet Res 2014 Jan;16(2):e38 [FREE Full text] [doi: 10.2196/jmir.3042] [Medline: 24513507]
3. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014 Jan;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]
4. 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]

Abbreviations

MeSH: Medical Subject Headings

XSL·FO
RenderX

Letter to the Editor

Response to: Self-Directed Interventions to Promote Weight Loss: A Systematic Review of Reviews

Melinda J Hutchesson[1,2], PhD; Philip J Morgan[2,3], PhD; Penelope McCoy[4], BHSc(N&D); Clare E Collins[1,2], PhD

[1]School of Health Sciences, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[2]Priority Research Centre in Physical Activity and Nutrition, University of Newcastle, Callaghan, Australia

[3]School of Education, Faculty of Education and Arts, University of Newcastle, Callaghan, Australia

[4]SP Health Pty Ltd, North Sydney, Australia

Corresponding Author:
Melinda J Hutchesson, PhD
School of Health Sciences
Faculty of Health and Medicine
University of Newcastle
Hunter Building HA12 University Drive
Callaghan, 2298
Australia
Phone: 61 49215405
Fax: 61 2 49217053
Email: melinda.hutchesson@newcastle.edu.au

Related Article:

Comment on: http://www.jmir.org/2014/2/e58/

In the February issue of the *Journal of Medical Internet Research*, Tang et al reported a systematic review of reviews examining the effectiveness of self-directed interventions to promote weight loss [1]. They reviewed 20 systematic reviews which incorporated 99 primary studies relevant to their review question. They concluded that self-directed interventions promote weight loss both independently and when provided as an adjunct to personal contact interventions.

We agree that the evidence presented in this systematic review of reviews provides some support for the use of self-directed interventions for weight loss; however we believe the data provided does not specifically support the use of self-directed interventions alone and when used in combination with other delivery modes. We wish to point out that this is because the majority of reviews included do not have inclusion criteria or present results in a way that considers whether the intervention was delivered solely in a self-directed format or in combination with other delivery modes (eg, face-to-face). In most of the reviews, these two types of self-directed interventions are grouped together when results are presented in a narrative summary or meta-analysis. We believe that as part of their recommendation for a comprehensive review of primary evaluations, the authors should have recommended that future systematic reviews of self-directed interventions should consider results separately for interventions that are delivered *solely* using self-directed modes, and those that use self-directed modes *in*

combination with other delivery modes. Furthermore, we believe that authors should be encouraged to describe the components of self-directed modes in sufficient detail to enable such a classification.

We are also concerned about the accuracy of the methodological quality review of Tang et al. For example, the review incorrectly states that Neve et al [2] "did not use quality assessment". However, the methods section of this review reports that "included studies were assessed by two independent reviewers for methodological validity using a standardized critical appraisal instrument from the JBI Meta-Analysis of Statistics Assessment and Review Instrument" (p 307) and the tool is provided as an Appendix to the paper [2]. Furthermore, the results include a description of the critical appraisal findings (subheading "Quality of Included Studies" p 310)[2]. As the results of the review of methodological quality are only presented as an overall score, and not by the individual criteria, it is difficult to determine how this influences the results of the quality evaluation within the review by Tang et al [1].

We agree with the authors that meta-analyses of self-directed weight loss interventions should be undertaken to help identify what intervention techniques and delivery formats are most effective. However, as an author of one of the included systematic reviews, we are aware that the absence of such meta-analyses from the included reviews is likely due to

inadequate descriptions of intervention techniques and delivery formats currently being published. Therefore, we believe a more important recommendation is that researchers who are currently evaluating self-directed interventions should provide more robust descriptions of their interventions, as recommended by the CONSORT-EHEALTH statement [3]. Furthermore, this may be achieved by specifically publishing study protocols (eg, [4,5]), or descriptions of the intervention development methods (eg, [6]).

References

1. Tang J, Abraham C, Greaves C, Yates T. Self-directed interventions to promote weight loss: a systematic review of reviews. J Med Internet Res 2014;16(2):e58 [FREE Full text] [doi: 10.2196/jmir.2857] [Medline: 24554464]
2. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of Web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
3. 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]
4. JMIR Research Protocols. E-collection: Randomized Controlled Trials - Protocols/Proposals (eHealth) URL: http://www. researchprotocols.org/themes/87 [accessed 2014-07-24] [WebCite Cache ID 6RJGbZ4JI]
5. JMIR Research Protocols. E-collection: Randomized Controlled Trials - Protocols/Proposals (funded, already peer-reviewed) URL: http://www.researchprotocols.org/themes/81 [accessed 2014-07-24] [WebCite Cache ID 6RJGZGIoQ]
6. JMIR Research Protocols. E-collection: Formative Studies and eHealth/mHealth Development URL: http://www. researchprotocols.org/themes/86 [WebCite Cache ID 6RJHPJEkb]

XSL•FO
RenderX

Virtual Communities of Practice: Overcoming Barriers of Time and Technology

Kieran Walsh[1*]; Stephen Barnett[2*], BMed, DCH(Lond), MRCGP, FRACGP

[1]BMJ Learning, London, United Kingdom

[2]General Practice Academic Unit, School of Science Medicine and Health, University of Wollongong, Wollongong, Australia

[*]all authors contributed equally

Corresponding Author:
Kieran Walsh
BMJ Learning
Tavistock Square
London, WC1H 9JR
United Kingdom
Phone: 44 7985755333
Fax: 44 02073836242
Email: kmwalsh@bmj.com

Related Article:

Comment on: http://www.jmir.org/2014/3/e83/

Letter

Barnett et al should be congratulated for their fascinating report into the use of virtual communities of practice for family physician training [1]. It is interesting that the respondents reported time as a barrier to their participation in the community of practice. Undoubtedly, it would have taken them time to take part, but the amount of time needed should be placed in context and compared with the time needed to partake in face-to-face communities of practice. Face-to-face participation would take far more time—not least the time required to travel to the venue of the community of practice meeting. So in truth, virtual communities of practice are far more likely to be time saving. However the feeling that time was a barrier cannot be dismissed out of hand. The fact that trainees felt it to be a barrier may say more about how the medical education community views the virtual world. Do we take learning in virtual communities as seriously as we take face-to-face learning? What would be the response of a tutor or trainer if a trainee requested study leave to participate in online education? Would they be as accommodating as they would be to a trainee who requested leave to attend an event? Is virtual learning something that trainers and trainees see as an activity to be undertaken in users' own time? These questions are important—not least because time is money in all areas of life and not least in medical education. [2]

Second, there was the issue of technology and some users feeling less comfortable using chats and wondering whether webinars might be made available in future generations of the virtual community. With the increasing availability of broadband Internet and the players required for creating and watching webinars, there is no reason why this cannot happen. The users may also discover unexpected benefits from webinars. Foremost among these might be the ability to learn communication and team-working skills, which they will be able to do much more easily by means of watching and interacting with videos. This might be particularly useful for users who have not trained in Australia.

Kieran Walsh

Author's Response

Dr Walsh makes two very good points in his letter; firstly, that the perception of online time needs to be offset against the time taken for face-to-face training and secondly, that webinars are a viable option for online training.

Regarding the perception of time, in a training context this fits well with the findings from the ConnectGPR study that it is important that the training is recognized, rather than just being an "extra". If the training time is recognized so that it offsets

face-to-face time, then perhaps that perceived time barrier will diminish.

In terms of webinars, in fact the ConnectGPR project is continuing in Australia, and for the last 12 months we have been running regular webinars, rather than webchats, which have been very well received. This year, our project is to try to supply enough webinar training that is live and then available later as a recording, that registrars will be able to offset face-to-face

workshop time by participating in these online events. Whilst face-to-face workshops are important and will never be completely replaced, for a training program that covers 160,000 square kilometres, there are potentially significant savings to be made in workshop attendance if even some of the face-to-face time can be replaced with online training.

Stephen Barnett

References

1. Barnett S, Jones SC, Caton T, Iverson D, Bennett S, Robinson L. Implementing a virtual community of practice for family physician training: a mixed-methods case study. J Med Internet Res 2014;16(3):e83 [FREE Full text] [doi: 10.2196/jmir.3083] [Medline: 24622292]
2. Walsh K. Medical education: time is money. J Surg Educ 2014 Apr;71(2):278-279. [doi: 10.1016/j.jsurg.2013.08.006] [Medline: 24602720]

XSL•FO

RenderX

Corrigenda and Addenda

Metadata Correction: How Patients With Schizophrenia Use the Internet: Qualitative Study

Beate Schrank[1], MD, MSc; Ingrid Sibitz[1], MD; Annemarie Unger[1], MD; Michaela Amering[1], MD

Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria

Corresponding Author:
Beate Schrank, MD, MSc
Department of Psychiatry and Psychotherapy
Medical University of Vienna
Waehringeguertel 18-20
Vienna, 1090
Austria
Phone: 43 1 40400 3603
Fax: 43 1 40400 3605
Email: beate.schrank@gmail.com

Related Article:

Correction of: http://www.jmir.org/2010/5/e70/

The authors of "How Patients With Schizophrenia Use the Internet: Qualitative Study" (J Med Internet Res 2010 Oct-Dec; 12(5): e70) have inadvertently declared on submission that all authors contributed equally. The author Schrank should have been listed as the sole first author, with the authors Sibitz, Unger, and Amering listed as co-authors, with adequate but not equal contributions. All authors have agreed to remove the "equal contribution" footnote (see Multimedia Appendix). Consequently, the "equal contribution" footnote has been removed in the corrected version of the paper on the JMIR website together with publishing this correction notice on July 30, 2014. A correction notice has been sent to PubMed and the correct full-text has been resubmitted to Pubmed Central and other full-text repositories.

Multimedia Appendix 1

Statement signed by all authors agreeing to the corrigendum of authorship contributions.

[PDF File (Adobe PDF File), 1MB - jmir_v16i7e165_app1.pdf]

XSL•FO
RenderX