

Journal of Medical Internet Research

Impact Factor (2018): 4.945 - ranked #1 medical informatics journal by Impact Factor
Volume 14 (2012), Issue 3 ISSN: 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Reviews

- Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review ([e80](#))
Catriona Kennedy, John Powell, Thomas Payne, John Ainsworth, Alan Boyd, Iain Buchan. 4
- Health and Wellness Technology Use by Historically Underserved Health Consumers: Systematic Review ([e78](#))
Enid Montague, Jennifer Perchonok. 183

Editorial

- Beyond Efficacy: The Depth and Diversity of Current Internet Interventions ([e92](#))
Helen Christensen, Alison Caele, Gerhard Andersson, Frances Thorndike, Robert Tait. 24

Original Papers

- Effectiveness of a Web-Based Multiple Tailored Smoking Cessation Program: A Randomized Controlled Trial Among Dutch Adult Smokers ([e82](#))
Eline Smit, Hein de Vries, Cisca Hoving. 28
- Efficacy of a Computerized Simulation in Promoting Walking in Individuals With Diabetes ([e71](#))
Bryan Gibson, Robin Marcus, Nancy Staggers, Jason Jones, Matthew Samore, Charlene Weir. 40
- Short-term Effects of a Smoking Prevention Website in American Indian Youth ([e81](#))
Deborah Bowen, Patricia Henderson, Jessica Harvill, Dedra Buchwald. 55
- Design of an mHealth App for the Self-management of Adolescent Type 1 Diabetes: A Pilot Study ([e70](#))
Joseph Cafazzo, Mark Casselman, Nathaniel Hamming, Debra Katzman, Mark Palmert. 63
- There's an App for That: Content Analysis of Paid Health and Fitness Apps ([e72](#))
Joshua West, P. Hall, Carl Hanson, Michael Barnes, Christophe Giraud-Carrier, James Barrett. 77
- Researchers' Needs for Resource Discovery and Collaboration Tools: A Qualitative Investigation of Translational Scientists ([e75](#))
Suresh Bhavnani, Michael Warden, Kai Zheng, Mary Hill, Brian Athey. 89

Mining Online Social Network Data for Biomedical Research: A Comparison of Clinicians' and Patients' Perceptions About Amyotrophic Lateral Sclerosis Treatments (e90)	
Carlos Nakamura, Mark Bromberg, Shivani Bhargava, Paul Wicks, Qing Zeng-Treitler.	99
Use of Social Media by Western European Hospitals: Longitudinal Study (e61)	
Tom Van de Belt, Sivera Berben, Melvin Samsom, Lucien Engelen, Lisette Schoonhoven.	106
Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers (e91)	
Liesje Donkin, Nick Glozier.	115
Effectiveness of an Online Group Course for Depression in Adolescents and Young Adults: A Randomized Trial (e86)	
Rianne van der Zanden, Jeannet Kramer, Rob Gerrits, Pim Cuijpers.	127
A Novel Evaluation of World No Tobacco Day in Latin America (e77)	
John Ayers, Benjamin Althouse, Jon-Patrick Allem, Daniel Ford, Kurt Ribisl, Joanna Cohen.	141
Accuracy of Geographically Targeted Internet Advertisements on Google Adwords for Recruitment in a Randomized Trial (e84)	
Ray Jones, Lesley Goldsmith, Christopher Williams, Maged Kamel Boulos.	152
Interaction Patterns of Nurturant Support Exchanged in Online Health Social Networking (e54)	
Katherine Chuang, Christopher Yang.	205
Building a Transnational Biosurveillance Network Using Semantic Web Technologies: Requirements, Design, and Preliminary Evaluation (e73)	
Douglas Teodoro, Emilie Pasche, Julien Gobeill, Stéphane Emonet, Patrick Ruch, Christian Lovis.	226
Using Internet Search Engines to Obtain Medical Information: A Comparative Study (e74)	
Liupu Wang, Juexin Wang, Michael Wang, Yong Li, Yanchun Liang, Dong Xu.	243
Sensitivity and Predictive Value of 15 PubMed Search Strategies to Answer Clinical Questions Rated Against Full Systematic Reviews (e85)	
Thomas Agoritsas, Arnaud Merglen, Delphine Courvoisier, Christophe Combescure, Nicolas Garin, Arnaud Perrier, Thomas Perneger.	262
What do Web-Use Skill Differences Imply for Online Health Information Searches? (e87)	
Markus Feufel, S Stahl.	277
Parent Opinions About Use of Text Messaging for Immunization Reminders (e83)	
Carolyn Ahlers-Schmidt, Amy Chesser, Angelia Paschal, Traci Hart, Katherine Williams, Beryl Yaghmai, Sapna Shah-Haque.	288
Disseminating Self-Help: Positive Psychology Exercises in an Online Trial (e63)	
Stephen Schueller, Acacia Parks.	297
Self-monitoring Using Mobile Phones in the Early Stages of Adolescent Depression: Randomized Controlled Trial (e67)	
Sylvia Kauer, Sophie Reid, Alexander Crooke, Angela Khor, Stephen Hearps, Anthony Jorm, Lena Sanci, George Patton.	308
Individually Tailored Internet-Based Treatment for Young Adults and Adults With Panic Attacks: Randomized Controlled Trial (e65)	
Kristin Silfvernagel, Per Carlbring, Julia Kabo, Sara Edström, Jenny Eriksson, Lisa Månson, Gerhard Andersson.	325

Superusers in Social Networks for Smoking Cessation: Analysis of Demographic Characteristics and Posting Behavior From the Canadian Cancer Society's Smokers' Helpline Online and StopSmokingCenter.net (e66)	
Trevor van Mierlo, Sabrina Voci, Sharon Lee, Rachel Fournier, Peter Selby.	338
Web-Based Cognitive Behavior Therapy for Depression With and Without Telephone Tracking in a National Helpline: Secondary Outcomes From a Randomized Controlled Trial (e68)	
Louise Farrer, Helen Christensen, Kathleen Griffiths, Andrew Mackinnon.	353
From Online Randomized Controlled Trials to Participant Preference Studies: Morphing the San Francisco Stop Smoking Site into a Worldwide Smoking Cessation Resource (e64)	
Ricardo Muñoz, Adrian Aguilera, Stephen Schueller, Yan Leykin, Eliseo Pérez-Stable.	363
Working Alliance in Online Cognitive Behavior Therapy for Anxiety Disorders in Youth: Comparison With Clinic Delivery and its Role in Predicting Outcome (e88)	
Renee Anderson, Susan Spence, Caroline Donovan, Sonja March, Samantha Prosser, Justin Kenardy.	375
Internet-Based Interventions to Promote Mental Health Help-Seeking in Elite Athletes: An Exploratory Randomized Controlled Trial (e69)	
Amelia Gulliver, Kathleen Griffiths, Helen Christensen, Andrew Mackinnon, Alison Caley, Alison Parsons, Kylie Bennett, Philip Batterham, Rosanna Stanimirovic.	410

Short Paper

Clinicians' Perspectives on a Web-Based System for Routine Outcome Monitoring in Old-Age Psychiatry in the Netherlands (e76)	
Marjolein Veerbeek, Richard Oude Voshaar, Anne Pot.	167

Tutorial

Using Crowdsourcing Technology for Testing Multilingual Public Health Promotion Materials (e79)	
Anne Turner, Katrin Kirchhoff, Daniel Capurro.	174

Viewpoint

Using Instructional Design Process to Improve Design and Development of Internet Interventions (e89)	
Michelle Hilgart, Lee Ritterband, Frances Thorndike, Mable Kinzie.	392

Letters to the Editor

From eHealth Technologies to Interventions (e93)	
Rik Crutzen.	428
Consistently Increasing Numbers of Online Ratings of Healthcare in England (e94)	
Felix Greaves, Christopher Millett.	431

Review

Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review

Catriona M Kennedy^{1,2}, PhD; John Powell³, PhD; Thomas H Payne⁴, MD, FACP; John Ainsworth¹, MSc; Alan Boyd⁵, MSc; Iain Buchan¹, MD, FFPH

¹School of Community-Based Medicine, University of Manchester, Manchester, United Kingdom

²School of Computer Science, University of Birmingham, Birmingham, United Kingdom

³Division of Health Sciences, University of Warwick, Warwick, United Kingdom

⁴UW Medicine, University of Washington, Seattle, WA, United States

⁵Health Policy and Management Group, Manchester Business School, Manchester, United Kingdom

Corresponding Author:

Iain Buchan, MD, FFPH

School of Community-Based Medicine

University of Manchester

1.311 Jean McFarlane Building

Oxford Road

Manchester, M13 9PL

United Kingdom

Phone: 44 161 275 5205

Fax: 44 161 275 5160

Email: Buchan@manchester.ac.uk

Abstract

Background: Information technology can help individuals to change their health behaviors. This is due to its potential for dynamic and unbiased information processing enabling users to monitor their own progress and be informed about risks and opportunities specific to evolving contexts and motivations. However, in many behavior change interventions, information technology is underused by treating it as a passive medium focused on efficient transmission of information and a positive user experience.

Objective: To conduct an interdisciplinary literature review to determine the extent to which the active technological capabilities of dynamic and adaptive information processing are being applied in behavior change interventions and to identify their role in these interventions.

Methods: We defined key categories of active technology such as semantic information processing, pattern recognition, and adaptation. We conducted the literature search using keywords derived from the categories and included studies that indicated a significant role for an active technology in health-related behavior change. In the data extraction, we looked specifically for the following technology roles: (1) dynamic adaptive tailoring of messages depending on context, (2) interactive education, (3) support for client self-monitoring of behavior change progress, and (4) novel ways in which interventions are grounded in behavior change theories using active technology.

Results: The search returned 228 potentially relevant articles, of which 41 satisfied the inclusion criteria. We found that significant research was focused on dialog systems, embodied conversational agents, and activity recognition. The most covered health topic was physical activity. The majority of the studies were early-stage research. Only 6 were randomized controlled trials, of which 4 were positive for behavior change and 5 were positive for acceptability. Empathy and relational behavior were significant research themes in dialog systems for behavior change, with many pilot studies showing a preference for those features. We found few studies that focused on interactive education (3 studies) and self-monitoring (2 studies). Some recent research is emerging in dynamic tailoring (15 studies) and theoretically grounded ontologies for automated semantic processing (4 studies).

Conclusions: The potential capabilities and risks of active assistance technologies are not being fully explored in most current behavior change research. Designers of health behavior interventions need to consider the relevant informatics methods and algorithms more fully. There is also a need to analyze the possibilities that can result from interaction between different technology

components. This requires deep interdisciplinary collaboration, for example, between health psychology, computer science, health informatics, cognitive science, and educational methodology.

(*J Med Internet Res* 2012;14(3):e80) doi:[10.2196/jmir.1893](https://doi.org/10.2196/jmir.1893)

KEYWORDS

Behavior change; consumer health informatics; health communication; health promotion; personalization

Introduction

Prevention, early intervention, and self-care are priorities for most health care systems around the world. Policy makers cannot, however, address these priorities solely through conventional clinical means. This is because citizens must make sustained health behavior changes, which are largely beyond the reach of the clinic [1].

Information technology has the potential to support behavior change [2]. However, in many behavior change interventions, technology is used as only a passive medium, where the main purpose is efficiency, communication, or a positive user experience. For example, an interactive website might provide different types of multimedia or a discussion forum, but the health information may be static. Even if the health content is tailored to a particular audience, the tailoring is determined a priori. Users navigate the website in the same way as they would navigate tailored print material, albeit often with greater ease. Processing of health content is not automated.

Furthermore, many interventional studies do not identify which *aspects* of technology contributed to the observed outcomes; the technologies are simply named in a general way (eg, mobile phones [3]). Recent reviews on technology-supported interventions [4-7] also did not explore the semantic processing potential of the technology. For example, Enwald and Huotari [5] defined “second generation” as “interactive media” and “third generation” as the use of “mobile and remote devices.”

In this paper we aim to review research on the *active* aspects of technologies to assist health behavior change. Our objectives were (1) to define key technological capabilities involving active information processing, (2) to determine through a literature review where these technologies are used in behavior change interventions or where they are clearly recognized as important for interventions, (3) to identify emerging themes involving novel uses of active technology that are cross-disciplinary, and (4) to identify any gaps in the research and propose future directions.

Definition: Active Assistance

We define active assistance technology as any technology involving *automated processing of health or behavior change information that is ongoing as the user interacts with the technology*. In other words, the technology continues to process the health-related information during interactive use and may adapt its responses. This contrasts with passive information technology such as storage devices, videos, and website design. It also contrasts with interactive systems that do not process information related to health or behavior change. For example, an interactive system can process user choices on presentation format (eg, video or text) and adapt to these choices during a

session. However, this is not active assistance because the responses are not related to the semantic content of the health messages, only to their formatting.

In this way, the concept of active assistance draws attention to the distinction between semantic and nonsemantic information processing during an interactive session. This is important, because semantic processing entails a degree of delegation of health decision making to an automated system, which can free up human specialists. It also has more serious consequences if incorrect.

Furthermore, active assistance takes place in an environment in which citizens and experts participate actively in the behavior change intervention (combining push and pull). In other words, the role of the technology is not merely to deliver a fully expert-led intervention where users follow instructions with minimal understanding. Instead, the technology helps users to reflect and learn about the obstacles to successful behavior change. A desirable feature is that users should feel they have ownership or control of the intervention [8].

A concept related to active assistance is persuasive technology, which is “designed to change attitudes or behaviors of users through persuasion and social influence” [9]. Although many aspects of persuasion are relevant to active assistance, persuasive technology is not necessarily active in the way that we are defining it; passive techniques such as visualization and website design may suffice. However, some aspects of persuasion would benefit from active technology. In particular, the potential of the technology for unbiased information processing can help with self-monitoring and correction of inaccurate health beliefs, both of which are required in the persuasion model [10,11]. Similarly, many behavior change theories such as the transtheoretical model [12], social cognitive theory [13-15], and theory of planned behavior [16] imply the need for health education, accurate beliefs, and self-monitoring to effect behavior change and sustain it over the long term. Independently of persuasion and self-monitoring, automated information processing can also overcome some of the problems of self-report in measuring behavior change such as adherence [17].

In the context of these requirements, the following are key examples of active assistance technology that can support behavior change. The technologies may be used together or independently.

Automated Reasoning Using an Explicit Knowledge Representation

An example of this is that an interactive system might use knowledge-based reasoning about the user's health and circumstances to determine how its responses should be tailored

further to the particular individual. Since this process is dynamic, there is more potential for delivery of messages that are tailored to the user's current environment and state of motivation than would be the case for static tailoring. Similarly, health education can involve answering specific health-related questions, on demand, using inference about what-if scenarios that the user wants to know about (nor necessarily personalized). Examples might include mobile or Web interfaces with dynamic personalization, intelligent reminding, natural-language dialog, or health-related games with an automated player.

Automated Data Collection With Pattern Recognition (Smart Sensing)

Automated sensing can overcome the limits of self-reported online diaries, which depend on memory. Recognition of patterns in online interactions, physical activities, or physiological states can provide useful self-monitoring information for users who are attempting to change their behavior, provided that it is displayed in a user-friendly way. This goes beyond the capability of automated reasoning because the system can acquire data and recognize events automatically without manual data input.

Automated Adaptation Over Time

Adaptation occurs in response to emerging patterns and contexts. This goes beyond automated reasoning and automated data collection because the system adapts its methods and decisions according to patterns that it has recognized. For example, an interactive system might learn to predict a user's state of motivation based on his or her responses to prompts (without any additional sensing).

In each case, the algorithms need to be informed by health-related knowledge, either explicitly as a formal representation or implicitly in the form of assumptions built into their design.

Background: Related Reviews

Some previous literature reviews have addressed topics related to active assistance.

Webb et al [2] conducted a systematic review and meta-analysis of the effectiveness of Internet-based interventions. The review found that interventions grounded in behavior change theories are associated with increased effectiveness, with the theory of planned behavior having the largest positive effect. The review used a taxonomy of behavior change techniques and found that using a combination of these techniques tended to be more effective than using only one. Effectiveness was also increased by communication methods such as instant messaging. Although the review did not consider the role of active semantic processing, the findings that a combination of techniques is more effective suggest an increased necessity for such automated processing due to the added complexity of interventions that combine more than one technique.

Fry and Neff [18] reviewed the effectiveness of interventions using automated prompts and reminders, finding that more than half of those (11 out of 19 reviewed) showed positive results. The findings suggest that tailoring of prompts and making them more interactive can make them more effective. However, the

review did not specify the role of the technology in determining the timing and content of prompts: whether these are manually designed in advance, or whether they can be adapted using automated methods.

Lustria et al [4] defined second-generation interventions as those that are tailored to individuals. The definition did not specify whether the tailoring is an ongoing automated adaptation process that happens during continued user interaction. However, the review has some interesting findings. Many tailored interventions have components for developing self-regulatory skills and self-monitoring. For example, the user can often keep online diaries and send them to an expert for analysis, who then provides feedback. Automation of some of the semantic analysis steps in this process would reduce the need for human expert involvement, thereby reducing costs. Similarly, Revere and Dunbar [19] conducted an earlier review (1996–1999), finding that tailored interventions tended to be effective but not enough evidence that tailoring was superior to generic interventions. The review mentioned the possibility of dynamic tailoring but did not focus on the distinction between dynamic and offline computer tailoring.

Bickmore and Giorgino [20] reviewed the technologies used in dialog systems and presented some examples that have been evaluated as effective. Similarly, Corkrey and Parkinson [21] reviewed interactive voice response systems from 1989 to 2000, finding a growing research area but few applications at that time.

For the purpose of building personalized health models, Fernandez-Luque et al [22] reviewed the research on extracting content from social networking sites (eg, blogs, forums, or search patterns). These techniques have relevance in behavior change interventions but are still at an early stage and present many ethical and technical challenges.

There have been no comprehensive reviews of active assistance technologies in health-related behavior change.

Methods

We conducted the literature review in accordance with the guidelines of the PRISMA statement [23] that are relevant to our objectives.

Keyword Search and Databases

We used the following strategy: <technology-related keywords> AND <health psychology-related keywords> AND health AND "behavior change".

Technology- and psychology-related keywords were combinations of the following: [automated OR technology OR Internet OR "mobile phone" OR intelligent OR computer-based OR interactive OR agent-based OR adaptive OR "context-aware" OR "machine learning" OR "pattern recognition" OR robotic OR "virtual reality" OR semantic OR "knowledge-based" OR "decision support" OR ontology OR dialog OR "natural language"] AND [assistance OR intervention OR personalization OR persuasion OR adherence OR compliance OR motivation OR affective OR emotion OR reminder OR prompt].

We conducted some preliminary searches on a wide range of databases including CINAHL, EMBASE, Inspec, ISI Web of Science, PsycINFO, and ScienceDirect. However, we found that Google Scholar had a wide enough coverage to allow it be used instead of these databases. This is consistent with recent empirical studies such as those of Howland et al [24] and Walters [25]. Therefore, we decided to use only Google Scholar and PubMed. The date range was January 2005 to January 2012. Only articles written in English were included.

Inclusion and Exclusion Criteria

We used the definition of active assistance above. The studies should have included at least one of the three active assistance technologies listed and have been intended for interactive use by clients or patients attempting to change their behavior. Studies may have described a new technology or design to be used in behavior change interventions, or they may have reported an evaluation of an actual intervention using the technology. Qualitative studies are necessary to evaluate usability and acceptability of active technology. Similarly, prototypes and works in progress help to provide an overview of the current research concepts and their maturity.

We excluded the following kinds of study: (1) interventions merely delivered using the Internet, CD-ROM, or other medium, where the technology only facilitates transmission of information from expert to user, (2) feedback, in which there is no automated processing of personal health-related information—for example, receiving emails from a human counselor would be excluded, (3) tailored-offline interventions, in which the computer processing is used by health professionals to tailor the intervention before or after the user interacts with the technology—this is the case where the semantic content processing is not part of the interactivity with the client (although a health professional may have interactive access), (4) simple data collection or preprogrammed reminding without any pattern recognition or inference (eg, pedometers); interventions where the semantic content of reminder messages were configured by the client were also excluded for this reason—for example, the planning tool of Soureti et al [26] is a passive technology and therefore excluded, (5) fixed games or simulations, in which there is no automated player with semantic processing capability, and (6) general guidelines or research roadmaps in which different options are discussed.

Data Extraction

We divided studies into the following categories: (1) quantitative and qualitative evaluations of interventions using active technology, (2) pilot studies of new technology, and (3) prototypes or designs that are being developed or tested.

For each study, we asked the following questions. What kind of active technology was used? Was it effective? What was the role of the technology in the intervention? Was it theoretically grounded?

Active Technology Types

We looked for one or more of the following types of automated content processing, based on capabilities of the active assistance technologies outlined above: (1) automated data collection with

pattern recognition, (2) context-sensitive alerts, reminders, and recommendations, (3) knowledge-based reasoning or inference (semantic representation, ontology, decision support, decision algorithm, and automated planning), (4) dialog systems with natural-language processing, (5) simulation or game with an intelligent agent, and (6) online adaptation to build user models and personalization (adaptive websites or interfaces, and user profiling).

In addition to our predefined categories, we identified new technology themes and author keywords describing the technology.

Effectiveness Evaluations

For those studies with evaluations of effectiveness, we asked the following questions. First, what was being evaluated? This could be acceptability or usability (self-reported positive or negative attitude); treatment adherence or technology engagement (observed); self-reported behavior change; or objectively measured behavior change (eg, step counts). Second, what method of evaluation was used (eg, randomized controlled trial [RCT] or qualitative study)? Third, were findings summarized, to give an indication of the maturity of the technology, and any advantages or new problems that it introduces?

Role of Active Technology in Interventions

It is important to understand the role of the active technology in the intervention.

We used the following three functions (defined above): (1) dynamic tailoring, (2) interactive education to support participations in their own care and disease prevention, and (3) support for self-monitoring in a way that overcomes biases of self-report.

In particular, we looked for an association of an active technology type with a purpose. For example, pattern recognition and context awareness may be used to support dynamic tailoring. Similarly, for unbiased self-monitoring, the technology needs to provide automated data collection, pattern recognition, and representation of the results in a visual format that can be easily understood.

Theoretical Grounding

We included here any behavior change theories mentioned by the authors as having a role in the technology design. In addition, we asked whether the study proposed any novel ways of connecting active technology with behavior change theories, and whether the active technology allows new possibilities that would not be available with static technology.

Results

Following a review of title and abstracts, the search identified 228 potentially relevant articles. Of these, 41 satisfied the inclusion criteria after a full-text review. Table 1 lists the data extraction contents for intervention and active technology themes [27-67], along with any effectiveness evaluations. Table 2 [27-67] shows the extent to which our previously defined technology roles appeared in the studies (dynamic tailoring,

interactive education, and self-monitoring support). Also included are any behavior change theories informing the technology design or usage.

In Table 1, we used the following notation to give an approximate summary of evidence for studies with effectiveness evaluations: <weight of study>: <effect>. Weight of study was scored as 5 (RCT with at least one objective measure, long-term), 4 (RCT with at least one objective measure, short-term), 3 (RCT with self-report only, long-term), 2 (RCT

with self-report only, short-term), or 1 (qualitative or pilot study). Effect was scored as + (positive), (negative), or +/- (mixed or insignificant).

For example, a study with objective measures over the long term, but not showing a significant effect, would be summarized as 5: +/- . We used the same summarized notation if some measures were positive and others negative or insignificant. Details are in the findings column.

Table 1. Technology themes, study types, and main findings.

Reference and project or intervention name	Health topic / study population	Technology themes	Type of study	Main findings	Evidence summary (if applicable) ^a
Ananthanarayan & Siek 2010 [27] (HealthSense)	Obesity / children	Wearable computing, “6th sense;” actionable feedback.	Design of prototype to support children’s motivation for exercise and for self-monitoring.	Not an empirical study.	Not applicable
Arteaga et al 2009 [28]	Obesity / teenagers	Motivational agent (mobile phone games).	Design of prototype to motivate exercise based on personality type.	Not an empirical study.	Not applicable
Bickmore & Picard 2005 [29] (Fit-Track)	Physical activity / healthy adults	ECA ^b : relational agent.	RCT ^c (n = 101; 30 days). Measures: acceptability (self-report) + PA ^d (pedometer). Groups: relational agent, non-relational agent and control.	Positive acceptance; increased PA during intervention but reduced PA after follow-up. Relational agent more liked. Dialog repetitiveness annoying.	4: +/-
Bickmore et al 2005 [30] (Fit-Track)	PA / older adults	ECA: relational agent.	RCT (n = 21; 2 months), to test acceptability (usage history) + PA (pedometer) + loneliness (self-report). Groups: relational vs control (usual care).	Positive acceptance and significant increase in PA during intervention. No significant decrease in loneliness.	4: +/-
Bickmore & Sidner 2006 [31]	General behavior change / adults	Making dialog more robust by linking with ontologies for behavior change theories (TTM ^e , MI ^f).	Prototype.	Not an empirical study.	Not applicable
Bickmore et al 2009 [32]	Physical activity / adults (male students)	Context awareness of mobile PA monitor + ECA (relational agent).	Pilot study (n = 8): test whether agent context awareness promotes social bonding (acceptance). Effectiveness: does it promote walking?	Some positive acceptance but less actual walking in context-aware condition.	1: +/-
Bickmore et al 2009 [33]	Compliance / low health-literacy patients (hospital discharge)	ECA: virtual nurse with relational behaviors and empathy.	Self-report usability tests: 2 tests: nonpatients (n = 30) + patients (n = 19) with 47% low literacy. Both groups tested with relational vs non-relational agent.	Both tests: relational preferred. Overall ECA acceptance. ECA allows more time and sense of control than human face-to-face communication.	1: +
Bickmore et al 2010 [34]	Medication adherence, PA / schizophrenia patients	ECA: simple concrete communication. Authors counter ethical criticism of ECA for mental health.	Pilot evaluation (n = 20; 31 days) to test acceptability (self-report) + adherence + PA (no control).	Positive acceptance. Adherence + PA high. ECA provides simplified conversation, less confusing than human face-to-face.	1: +
Bickmore et al 2011 [35]	2 domains: exercise and diet / adults	Semantic ontology for behaviors and theories (TTM, MI); semantic models of user, data, and intervention.	Qualitative study (n = 8) on acceptability of ECA health counselor based on reusable ontology.	Positive acceptance, but limited evaluation.	1: +
Bickmore et al 2010 [36]	Physical activity / adults	ECA: promoting long-term use; avoid repetitive dialog. Introduce variability + storytelling.	2 RCTs: 1. Variability (n = 24, 100 days); variable vs nonvariable; 2. Story(n = 26, 30 days): first-person story vs third-person story. Measures: usage + step count + self-reported satisfaction.	1. Variability: more system usage, but less exercise. 2. Story: first person had more usage than third person, but less exercise. Self-reported satisfaction high for test conditions.	4: +/-
Bieber et al 2009 [37] (DiaTrace)	Physical activity / adults	Mobile phone as sensor for activities and calorie estimate.	Prototype.	Not an empirical study.	Not applicable

Reference and project or intervention name	Health topic / study population	Technology themes	Type of study	Main findings	Evidence summary (if applicable) ^a
Buttussi & Chittaro 2008 [38] (MOPET)	Physical activity / adults	ECA; context-aware sensing; user model.	Prototype.	Initial qualitative evaluation positive (n = 12).	1: +
Consolvo et al 2008 [39] (UbiFit)	Physical activity / adults	Graphic display with “garden” metaphor; mobile sensing device with inference; interactive app (edit or add to journal).	RCT: 3-month field experiment (n = 28): full system (10) vs no mobile sensing device (9) vs no display (9). Measures: (1) sensed activities + self-report; (2) qualitative analysis on user experience.	System with display led to more exercise than without display. User experience positive: more self-awareness, which motivated exercise.	4: +
De Rosis et al 2006 [40]	Diet / adults	ECA: recognize user’s emotional state, social attitude, and TTM stage during dialog; dynamically update user model during dialog.	Prototype: raters label emotional states, TTM stages, and social attitudes in test dialogs (WOZ ^g and corpus).	Labeling of emotions by raters used to guide design of dialog system.	Not an intervention evaluation
Farzanfar et al 2007 [41]	Treatment adherence, suicide prevention / depressed adults	Telephone agent: monitoring + self-care management.	Preliminary qualitative trial (n = 15), 4 weeks. Modifications made in response.	Dialog was helpful for adherence, but sounded artificial and insensitive, particularly in suicide risk. Users prefer more human-like agent with empathy and understanding of serious issues. (For suicide, hotline preferred). Authors’ conclusion: anthropomorphism is not valid (people do not attribute human qualities to machines—only in metaphor).	1: +/-
Hakulinen et al 2008 [42] (COMPANIONS project)	PA / adults	Mobile companion; semantic ontology of user environment for PA planning.	Prototype.	Not an empirical study.	Not applicable
Hartmann et al 2007 [43]	Improve patient questions to physician / adults with asthma	Educational website to suggest questions, encourage patient involvement in care, prevent more serious illness.	Pilot study: (n = 37) record usage experiences.	Positive self-report: (1) improved relations, (2) more active involvement.	1: +
Hayes et al 2009 [44]	Medication adherence / older adults	Instrumented pillbox, home sensors.	Pilot study (n = 10): effectiveness of context awareness on adherence. Test phases (same group): no-prompt, time-based, context-aware prompt.	Initial evaluation: positive for context-aware phase.	1: +
Jin 2010 [45]	Stress management / college students	Education-entertainment / health belief, self-efficacy; educational interactive test (game) for responses to stress scenarios. Agent gives educational messages.	RCT: (n = 60). Effectiveness of virtual agent on student’s intent/mood. Interactive test with agent (test) vs no agent (control) vs no test (true control).	Positive self-report on enjoyment and educational value for agent condition. Interactive test improves stress management self-efficacy (over true control, without test).	2: +

Reference and project or intervention name	Health topic / study population	Technology themes	Type of study	Main findings	Evidence summary (if applicable) ^a
Kaipainen et al 2011 [46] (Health-PGS)	General health decisions / adults	Health Personal Guidance System: guide user through day-to-day choices in ecosystem. Virtual individual: maintains user profile and context; HealthGuide: planning, context-aware messages. Personal Guidance System Mail: services all in one place.	Prototype.	Not an empirical study.	Not applicable
Klein et al 2011 [47] (eMate)	Adherence / diabetic patients	Automated reasoning based on COMBI ^h model (combines different theories) ensures dynamic tailored messages depending on user's context and state of mind.	Prototype: computational model of behavior change (mobile + website).	Not an empirical study.	Not applicable
Konovalov et al 2010 [48] (GATE)	Mental health promotion / military service personnel	Blog analysis to understand moods and emotions (combat experience): GATE algorithm + ontology.	Design and pilot study for technology: compare algorithm with expert opinion.	Precision of algorithm: 0.9, recall: 0.75; <i>F</i> score: 0.82.	Not an intervention evaluation.
Lee et al 2010 [49]	Health promotion / older adults	Telehealth: action-based behavior model (1) increase user's awareness of health, (2) set goals, (3) educate user in how to achieve goal, (4) remind, (5) reward + assess.	Design: overcome limits of sensing only; need high-level assessment information with models of persuasion to determine whether behavior changed.	Not an empirical study.	Not applicable
Levin & Levin 2006 [50]	Pain management / adults	Ecological momentary assessment, detect unexpected errors in dialog.	Feasibility study: evaluate interactive voice response system dialog for health and behavior monitoring. Feasibility study for pain monitoring voice diary (n = 24). 171 dialog sessions.	Accuracy of voice recognition: 98%. Dialog efficiency increased with user experience.	1: +
Lisetti & Wagner 2008 [51] (ABLE)	Mental health promotion / adolescent	ECA companions.	Design: ECA companion to act as MI counselor.	Not an empirical study.	Not applicable
Looije et al 2010 [52] (SuperAssist)	Adherence / older adults	ECA (robot cat), MI, persuasion.	Pilot study (n = 24): physical ECA (n = 12) vs virtual (n = 12). Each group experienced text, social ECA, and nonsocial ECA.	90% acceptance. Social ECA preferred over nonsocial ECA; half preferred text interface over social ECA ("conscientious" personality type). Virtual ECA more "empathic" than physical.	1: +
Maier et al 2010 [53] (SEMPER)	Work-related disorders and alcohol / adults	Semantic Web portal; semantic search.	Prototype.	Not an empirical study.	Not applicable
Mazzotta et al 2007 [54] (PORTIA)	Healthy eating / adults	Persuasion agent: tailoring of messages based on inferred personality traits and likely motivations of user.	Prototype of dialog design, based on corpus analysis of persuasive dialogs produced by participants in role-playing scenarios.	Corpus analysis found that persuasion is most often based on nonrational arguments and positive framing.	Not an intervention evaluation

Reference and project or intervention name	Health topic / study population	Technology themes	Type of study	Main findings	Evidence summary (if applicable) ^a
Munguia Tapia 2008 [55]	Obesity / adults	Sensors and algorithms for activity recognition and energy expenditure estimate.	Prototype.	Activity recognition most accurate if simple examples are given; high variability is difficult (eg, housework). Energy estimate more accurate for simple activities and with multiple body sensors.	Not an intervention evaluation
Nguyen & Mas-thoff 2008 [56]	General behavior change / adults	Persuasive dialog, MI.	Acceptability test (n = 41): is MI dialog more persuasive than argumentation? Questionnaire + qualitative analysis in comments.	Self-report positive: persuasiveness, likeability scores higher for MI than for 2 types of argumentation.	1: +
Oddsson et al 2009 [57] (SKOTEE)	Adherence / adults	Robotic assistance for intelligent reminding and companionship.	Design.	Not an empirical study.	Not applicable
Op den Akker et al 2011 [58]	PA / adults	Software agent for smart phone: use machine learning to develop user model. Tailor messages to user history and current context.	Prototype.	Not an empirical study.	Not applicable
Rojas-Barahona & Giorgino 2009 [59] (AdaRTE)	General behavior change / adults	Framework for health dialog.	Design.	Not an empirical study.	Not applicable
Smith et al 2008 [60]	Healthy lifestyle / adults	ECA; collaborative planning, update planned activities through ongoing dialog.	Prototype with technical evaluation.	Approach is feasible, although dialog error rate is still high.	Not applicable
Sorbi et al 2007 [61]	Migraine attack prevention / adults	PDA ⁱ + coaching. Response behaviors to precursors of migraine. Ecological momentary intervention experience sampling: randomized calls overcome memory bias. Tailored messages depending on current experience.	Pilot study (n = 5): feasibility and user acceptance.	Positive acceptance but too many calls are annoying. Technical problems: data loss due to buildings.	1: +/-
Spring et al 2010 [62] (Make Better Choices-MBC)	Obesity / adults	PDA: find optimal advice for exercise; goal thermometers; "in the moment" decision support/multiple theories, including self-regulation. Study design.	Study design.	Design of a trial only.	Not applicable
Tiwari et al 2011 [63]	Adherence / older adults	Robotic assistance, dialog.	Prototype development using grounded theory participatory design.	Emerging themes: usability, empowerment, collaboration, and safety: used as requirements for dialog design.	Not applicable
Turunen et al 2011 [64] (COMPANIONS project)	Health and fitness / adults	Home and mobile health and fitness companion.	Pilot study (n = 20): feasibility of complex dialogues in home and mobile scenarios.	System behaves robustly in realistic experimental scenarios, but word error rates are still high.	1 +
Uribe et al 2011 [65]	Adherence general	Reminders based on inferred mental state; user modeling using ontologies.	Prototype.	Not an empirical study.	Not applicable

Reference and project or intervention name	Health topic / study population	Technology themes	Type of study	Main findings	Evidence summary (if applicable) ^a
van der Putten et al 2011 [66] (SERA projet–Social Engagement with Robots and Agents)	PA / older adults	Social robot; health advisor.	Pilot study (n = 6). Video recording of interactions in homes. Iteratively modify setup based on results of previous session.	3 iterations, variable interactions, and satisfaction reports. Positive for motivation but some frustration over lack of control of dialog and too much time taken up.	1: +/-
Watson et al 2012 [67]	PA / overweight adults	ECA: virtual coach	RCT (n = 70; 12 weeks); primary measure: step count; secondary: weight + self-reported satisfaction, self-efficacy, PA recall, and PA stage of change. Groups: virtual coach vs control (no coach: website + pedometer only).	Average step count for intervention group remained constant over 12 weeks while control group dropped. Repeated measures analysis of variance showed significant difference in step count change between intervention and control. No significant difference in secondary measures; acceptance mixed.	4: +/-

^a <weight of study>: <effect>: weight of study was scored as 5 (randomized controlled trial [RCT] with at least one objective measure, long-term), 4 (RCT with at least one objective measure, short-term), 3 (RCT with self-report only, long-term), 2 (RCT with self-report only, short-term), or 1 (qualitative or pilot study). Effect was scored as + (positive), (negative), or +/- (mixed or insignificant).

^b Embodied conversational agent.

^c Randomized controlled trial.

^d Physical activity.

^e Transtheoretical model.

^f Motivational interviewing.

^g Wizard of Oz study, where humans pretend to be dialog agents to understand the likely responses to an automated agent.

^h Computerized behavior intervention.

ⁱ Personal digital assistant.

Table 2. Active technology role and theoretical grounding.

Reference and project or intervention name	Active technology type	Dynamic tailoring	Interactive education	Self-monitoring	Theoretical grounding
Ananthanarayan & Siek 2010 [27] (Health-Sense)	Inference; pattern recognition	Not specified	Yes, but details not given	Yes; provide awareness of physical activity	General awareness only; no specific theory mentioned
Arteaga et al 2009 [28]	Dialog; pattern recognition	Not specified. Static tailoring only	No	Not specified	Big 5 personality theory; technology acceptance model; theory of planned behavior, theory of meaning behavior
Bickmore & Picard 2005 [29] (FitTrack)	Dialog	Not specified	No; passive educational content only	Very basic, pedometer steps only	Relational agents
Bickmore et al 2005 [30] (FitTrack)	Dialog	Not specified	No; passive educational content only	Progress charts only	Relational agents
Bickmore & Sidner 2006 [31]	Inference; dialog	Not specified, but possible	No	Progress charts only	TTM ^a , MI ^b : link with agent reasoning and ontology
Bickmore et al 2009 [32]	Pattern recognition	Not specified in detail	No	No	Relational agents
Bickmore et al 2009 [33]	Dialog	Not specified in detail, only mentioned as a property of dialog systems in general	Yes, support low health-literacy patients	No	Relational agents
Bickmore et al 2010 [34]	Dialog	Not specified in detail	No	Not considered usable by schizophrenia patients	Relational agents
Bickmore et al 2011 [35]	Inference; dialog; user models	Not specified; fixed tailoring only mentioned	No, but mentioned in a generic way	No	TTM, MI encoded in ontology for agent reasoning and user model
Bickmore et al 2010 [36]	Dialog	Not specified	No	Charts only	Relational agents
Bieber et al 2009 [37] (DiaTrace)	Physical activity recognition, mobile phones	Not specified	No	No	Not mentioned
Buttussi & Chittaro 2008 [38] (MOPET)	Pattern recognition; adaptation; user model	Yes, due to context awareness	No	Not mentioned, but possible	Not mentioned
Consolvo et al 2008 [39] (UbiFit)	Activity recognition; inference	Not specified	No	Yes; visual display	No
De Rosis et al 2006 [40]	Dialog; user modeling, adaptation	Yes, due to adaptation	No	No	TTM
Farzanfar et al 2007 [41]	Dialog; pillbox sensors + adherence data analysis—linked to dialog system	Not specified, although possible	Yes, telephone instructions but limited interactivity	No	Self-efficacy theory, MI
Hakulinen et al 2008 [42] (COMPANIONS project)	Dialog; automated planning; knowledge-based inference	Not specified, although possible	No	No	Not mentioned
Hartmann et al 2007 [43]	Inference: evidence-based decision rules	Not specified in detail, but possible	Yes, but limited	No	No
Hayes et al 2009 [44]	Context-aware reminders; activity recognition; rule-based inference	Yes, decision to prompt based on recognized activity pattern	No	Not mentioned, but possible to include	Not mentioned
Jin 2010 [45]	Virtual agent in game	Not specified	Yes, education-entertainment	No	Health belief model, self-efficacy

Reference and project or intervention name	Active technology type	Dynamic tailoring	Interactive education	Self-monitoring	Theoretical grounding
Kaipainen et al 2011 [46]	Context awareness, pattern recognition, inference, planning, user modeling	Yes, messages tailored to changing context of user	Not a main focus	Not mentioned, but possible to include	Hybrid approach including self-efficacy and social influence
Klein et al 2011 [47] (eMate)	Knowledge-based reasoning; user models	Yes, automated reasoning based on COMBI ^c model ensures dynamic tailored messages depending on user's context and state of mind	No	No	COMBI model includes aspects of TTM, health belief model, social cognitive theory, self-regulation theories, attitude formation theory, and relapse prevention model; interaction based on MI
Konovalov et al 2010 [48]	Pattern recognition; inference	No, but could be used in an intervention with dynamic tailoring	No	No	No
Lee et al 2010 [49]	Pattern recognition; user modeling (profiling), including mental states	Not specified in detail, but planned	Not specified, but planned	Not specified, but planned	Action-based behavior model: (1) increase user's awareness of health; (2) set goals; (3) educate user in how to achieve goal; (4) remind; (5) reward + assess
Levin & Levin 2006 [50]	Voice recognition; semantic representation; dialog	Not specified, but personalization of dialog possible	No	No	No
Lisetti & Wagner 2008 [51] (ABLE)	Dialog system considered	Not specified, but possible	No	No	MI
Looije et al 2010 [52] (SuperAssist)	Dialog	Not specified, but possible	No	No	MI
Maier et al 2010 [53] (SEMPER)	Text mining; ontologies; machine learning; semantic search	Yes, personalized search results based on user profile built automatically	Yes, information portal	No	MI
Mazzotta et al 2007 [54] (PORTIA)	Dialog, user model	Yes, tailoring of persuasion messages based on inferred personality traits and likely motivations of user	No	No	Persuasion theories, argumentation
Munguia Tapia 2008 [55]	Activity recognition; energy estimate	No, but possible in an intervention	No	No, but possible in an intervention	No
Nguyen & Masthoff 2008 [56]	Dialog	Not specified	No	No	MI-based dialog design
Oddsson et al 2009 [57] (SKOTEE)	Intelligent reminding	Yes, part of robotic companion	No	Not mentioned, but possible to include	No
Op den Akker et al 2011 [58]	Pattern recognition, machine learning, context awareness, user modeling	Yes, messages are tailored based on user model and context	No	Not mentioned	No
Rojas-Barahona & Giorgino 2009 [59] (AdaRTE)	Dialog; adaptation	Yes, dialog can be adapted according to patient answers	No	No	No
Smith et al 2008 [60] (COMPANIONS)	Dialog control; inference; automated planning	Yes, update planned activities through ongoing dialog	No	No	No

Reference and project or intervention name	Active technology type	Dynamic tailoring	Interactive education	Self-monitoring	Theoretical grounding
Sorbi et al 2007 [61]	Adaptation, automated personalized feedback	Yes, tailored messages depending on current experience	No	No	No
Spring et al 2010 (Make Better Choices—MBC) [62]	Decision support; coaching algorithms. (PDA ^d)	Not specified, but possible	No	Yes, PDA allows this but not described in detail	No, although some theories mentioned
Tiwari et al 2011 [63]	Robot, dialog	Not specified in detail, but dynamic adaptation is a required feature in the design	No	No	No
Turunen et al 2011 [64] (COMPANIONS project)	Dialog; inference; automated panning	Yes, adaptive dialog, collaborative planning	No	No, but possible to include	No
Uribe et al 2011 [65]	Knowledge-based inference	Yes, reminders based on inferred mental state	No	Yes, implied in the design but not described in detail	TTM incorporated in ontology
van der Putten et al 2011 [66] (SERA project—Social Engagement with Robots and Agents)	Robot, dialog	Not mentioned	No	No	Not mentioned
Watson et al 2012 [67]	Dialog	Yes, dialog utterances tailored according to user progress with system	Not specified in detail	Not specified in detail	Relational agents

^a Transtheoretical model.

^b Motivational interviewing.

^c Computerized behavior intervention.

^d Personal digital assistant.

Emerging Technology Themes

From Table 1, it is clear that dialog systems were used widely in interventions using active technology (19 studies). Of these, embodied conversational agents (ECA) were components of 8 studies. ECAs are visual (*embodied*) characters that can conduct conversation with a user. They are agents in the sense that they can sense and respond to their environment [68]. The agent's environment might include a virtual game environment, the text inputs of a user, physical behavior data, or all three of these. Similarly, its responses might include conversational output, actions within the virtual environment, or effects on the real environment in the case of a robotic system [52,57,66]. Within the ECA context, motivation and empathy were central themes in 6 studies.

Ecological momentary assessment [69]) is a process of capturing the momentary experiences of participants—for example, using online diaries for the purposes of research. Not only are the participants' environmental and behavioral circumstances recorded but so are their mental states. These can include, for example, their current goals, beliefs, mood, and emotions. Ecological momentary interventions [70]) are based on ecological momentary assessment. Levin and Levin [50] conducted a feasibility study on pain monitoring and recommended management using interactive voice response.

Sorbi et al [61], in their pilot study, obtained a positive result for migraine attack prevention based on experience sampling using random mobile calls. Randomization overcomes memory bias. Another study that has relevance for the automation of ecological momentary assessment is the semantic analysis of blogs describing combat experience [48].

Effectiveness Evaluations

Most studies (18) were prototypes or design concepts. A total of 17 were feasibility or usability studies. Only 6 were RCTs measuring effectiveness for behavior change [29,30,36,39,45,67]. All of those also included a qualitative self-report of user experience. Bickmore and Picard [29] reported mixed results for an agent-based dialog system (FitTrack), with short-term improvement in physical activity during the intervention but reduced activity after follow-up. Interacting with an agent led initially to more physical activity than the control (nonagent) for sedentary participants (45 out of 101). Users enjoyed interacting with the relational agent more than with the nonrelational agent, but there was no significant difference in physical activity between the two agent conditions (both increased during the intervention). Additionally, users found the agent's repetitiveness to be annoying. The authors suggested that long-term interaction with the agent might prevent the drop in physical activity. Bickmore et al [30]

reported positive acceptance and increased physical activity for FitTrack with older adults ($n = 21$) over 2 months, but there was no significant decrease in loneliness.

The study of Bickmore et al [36] involved mechanisms for reducing the repetitiveness of a dialog system and enabling it to tell stories. The effect on long-term user engagement was positive, but the effect on actual behavior was negative. The authors proposed several explanations, such as increased enjoyment during interaction inhibiting activity. Their most recent research [67] reported a positive effect ($n = 70$) for sustained step count over 12 weeks, but there were mixed effects for satisfaction and motivation.

The study of Consolvo et al [39] involved a personalized display of the results of an activity recognition system and reported a positive effect on actual behavior and user experience (detailed below).

Jin [45] reported a positive effect of a virtual agent on stress management self-efficacy and enjoyment ($n = 60$).

Qualitative Studies

In qualitative pilot studies (17 studies), agents with empathy and social behavior tended to be preferred over nonsocial agents. In particular, Farzanfar et al [41] found that depressed adults needed more empathy and that they found the system to be insensitive in a suicide-risk situation. There were some results that were not predicted in advance. Bickmore et al [34] studied a system to support medication adherence for mental health patients. Participants found communication with the agent to be simpler than human face-to-face communication because they could slow down the interaction and give a greater sense of control. Bickmore et al [33] reported similar effects for low health-literacy patients.

Role of Active Technology

Dynamic Tailoring and User Modeling

A total of 15 studies emphasized dynamic tailoring. Of these, 10 were prototypes, 1 was an RCT [67] (detailed above), and 4 were pilot studies. The studies of Buttussi and Chittaro [38], Hayes et al [44], and Turunen et al [64] were positive but had limited evaluations. Sorbi et al [61] studied migraine prevention using ecological momentary interventions with positive acceptance, although too many calls could be annoying. Dynamic tailoring is particularly associated with user models that are generated or refined automatically. Common themes are context-aware activity recognition and intelligent reminding. Although all active assistance technology (as defined above) potentially has the capability for dynamic tailoring, we did not find this to be a major topic in most studies.

Interactive Education

Three studies [43,45,54] were directly concerned with health education. Hartmann et al [43] conducted a pilot study of a system to help people with asthma to participate in their own care by suggesting questions for them to ask their physicians using evidence-based decision rules. Jin's [45] RCT tested the effect of an agent within an educational game environment on

student stress management. The result was positive for self-reported stress management ($n = 60$).

Maier et al [53] developed a prototype for patient self-management for work-related disorders and alcohol reduction. One component was an information portal for training and health literacy, which was integrated with the Semantic Web.

Self-monitoring

Two studies on physical activity were concerned with accurate self-monitoring and visualization. Consolvo et al [39] conducted an RCT with positive results for a physical activity-awareness system for adults. The intervention (UbiFit) combined activity recognition with a visualization of the types of exercise that the system recognized. The visualization used a garden metaphor, where a particular type of flower represented a physical activity category (such as walking, cardiovascular, or strength). The user was awarded a flower when a physical activity was observed, eventually producing a varied garden (with a butterfly for completing the goal). Users had the opportunity to challenge and edit the system's inferences about their activities. Using the system made a positive difference to actual exercise behavior, with the visual display having a larger positive effect. The RCT involved 28 participants, with 10 assigned to the full system, 9 assigned to activity recognition only, and 9 assigned to a manual diary with display only.

Ananthanarayan and Siek [27] reported a design concept (HealthSense) to promote teenage physical activity, based on the principles of Consolvo et al [39] but aimed specifically at young people.

Theoretical Grounding

Behavior change models were used in 14 studies. Motivational interviewing [71] was the most widely used (8 studies). Of these, 5 were dialog systems, in which motivational interviewing was used as a general philosophy for dialog design [35,41,47,51,52]. Other behavior change theories that guided the designs were the transtheoretical model [31,35,40,47,65], self-efficacy [41,45], theory of planned behavior [28], and the health belief model [45].

Theoretically Grounded Ontologies

An important novel development in theoretically grounded active assistance is the incorporation of behavior change theories into the ontologies used in knowledge-based reasoning and dialog design (5 studies). The prototype in Bickmore and Sidner's study [31] and their follow-up pilot study [35] used ontologies for automated reasoning about the behavior change process. One of these ontologies is *theory-neutral*, while another is based on the transtheoretical model, enabling reasoning about the different stages. Similarly, De Rosis et al [40], Klein et al [47], and Uribe et al [65] incorporated aspects of the transtheoretical model and other theories into an ontology for dynamic tailoring of messages, depending on the inferred mental state of the user.

Discussion

The results show that significant research has been focused on dialog systems, ECAs, and activity recognition. There was also some work on ecological momentary intervention and intelligent context-aware prompting. The most covered health topic is physical activity. Most studies were still at an early stage, either prototype work in progress or pilot studies. Only 6 were RCTs, of which 4 were positive for behavior change and 5 were positive for acceptability.

The studies on dialog and ECA systems showed that empathy and relational behavior are significant research themes in behavior change, with many pilot studies showing preference for those features. The effect on actual behavior also tended to be positive. Too much interaction, however, might interrupt and inhibit positive health behaviors. So there is a need for careful consideration of the frequency and duration of interactive sessions in context.

Ecological momentary intervention is an opportunity for generating models from captured user experiences in the user's own language (eg, from social networking sites) and for integrating these models with expert knowledge. Such models can include the mental states and emotions of the user, particularly if they are used in conjunction with theoretically grounded ontologies [35]. Outside the behavior change domain, recent developments in mental health management have used models of this sort [72-74].

We found relatively few studies explicitly focusing on the functions of active technology that we selected above: dynamic tailoring, interactive education, and self-monitoring. Although some interventions may have included these functions implicitly, it seems that many studies did not recognize the role of a particular technology in enabling or improving these aspects.

Implications for Future Research

Links with Cognitive Science and Computer Science

Behavior change research needs to be informed by a deep understanding of algorithms and techniques that can support interventions. For this purpose, interdisciplinary collaboration with computer science and cognitive science is needed. In particular, behavior change technology has some parallels with educational technology. In educational systems, an intelligent tutor builds a model of the learner based on his or her performance and responses to questions (eg, what concepts does this person find difficult?) [75]. This model is then opened up for inspection so that the user can see how the system has represented his or her progress and misconceptions [76-80]. The open model supports the user's self-awareness, which is also an essential component of many behavior change theories. In behavior change research, we found only 1 study [45] that was aware of educational technology research.

Making users aware of the models can draw their attention to emotions and environmental circumstances (ecology) that are associated with negative behavior outcomes. Similarly, opening up models and giving users more control may enable users to spot any serious misunderstanding by an agent or dialog system,

thus avoiding the problem of users blindly following incorrect instructions. In some educational systems [81], users can persuade the automated tutor to change the model, because users are experts in their own experience (although not in the factual topics they are learning). This general principle of patient (or nonexpert) participation in health information and management is a current research topic in health informatics [82-86].

Alternatives to Dialog and ECA

Most studies on active assistance technologies in behavior change are based on natural-language dialog and ECA. We did not find many alternatives to these approaches that could be used if natural language or the ECA format is not suitable or preferable. For example, users might interact with adaptive interfaces where the users' actions are interpreted semantically as if they were dialog responses. Many of the core principles, such as model-based reasoning, activity recognition, and context-aware reminders, can be effective with different forms of interface.

Need for Dependable Systems

Studies on ECA and dialog systems are mostly focused on relational behavior and enjoyment of usage. If ECA systems are to be deployed in areas such as mental health and low health literacy [33,34] they will need to be validated as safe, effective, efficient, and acceptable to patients or clients. Such validation, for example in the European Union, may have to meet criteria usually applied to medical devices. This parsimonious approach makes it difficult to reflect realistic complexity—for example, the detection of emerging health problems based on subtle content of a conversation that a human expert would be able to detect. Sometimes, important decisions might be supported by ECA systems that fall outside of their validated uses. For example, errors might happen due to unexpected behavior of an algorithm. There is a need for research on making the systems robust in unplanned scenarios.

Need for Integrated Semantic Systems

Most studies in behavior change were focused on one or two technologies (eg, dialog and activity recognition) without specifying how the components can interact to infer further information. For example, coordination between activity recognition and content analysis of online diary entries might determine the circumstances in which relapses tend to occur, and tailor messages accordingly. Similarly, reliable automated decision making requires an interactive system to be connected with diverse specialist knowledge sources that can be requested on demand. More research is needed on how the components of an active assistance system are coordinated together and how they may be connected with the Semantic Web and other health informatics resources (eg, risk modeling).

Limits of This Review

Articles not indexed in Google Scholar or PubMed would have been missed—most scholarly publications, however, are captured by Google Scholar. The review required the mention of “health” and “behavior change” in the articles. We did not include gray literature such as white papers and unpublished reports. We selected the date range (2005–2012) to focus on

recent developments, but this may also have excluded innovative earlier work.

The review required specific mention of a key technology. There may be some interventions that use active technologies, but the studies did not mention this. Similarly, some studies mentioning only general intelligent technology were excluded from the full-text review because they could not be categorized. This may be a limitation because included studies need to involve significant interdisciplinary communication between technology specialists and health specialists. On the other hand, it may be a strength, as such communication is important for understanding a particular technology in context.

Since we limited the search to behavior change, it is also possible that many of the technologies are being applied in other areas of health informatics. For example, we found some prototypes early in the date range (2006–2007) but found no subsequent study relating to behavior change. In these cases, citation searching sometimes revealed further development of the techniques and algorithms, but no application in the health domain.

Conclusion

The potential of active technologies for dynamic and unbiased information processing is not being fully exploited in current health behavior change research. Most research has focused on specialist areas, such as dialog and ECA systems, and has been

largely restricted to the study of persuasive dialog in respect of relational behavior and motivation of behavior change.

In addition to the potential benefits of active technologies, there is a need for a thorough understanding of the potential risks. Expected benefits such as that of dynamic tailoring of the content and presentation of information can be measured using established evaluation methods (eg, [87,88]). Risks such as misinformation due to the unexpected behavior of an algorithm may be more difficult to uncover. It is important to study realistically complex scenarios of the uses of active assistance systems. Such studies need to reveal how the system components interact to produce information, and how these components might in turn interact with wider systems such as the Semantic Web, clinical records, and personal health records. Wider still, we note that many health behaviors are socially mediated; therefore, active assistance research needs to bridge cognitive and sociotechnical aspects in order to deliver maximum public health benefit.

To exploit the full potential of active assistance technology, health behavior change researchers need a deep understanding of how the different components of information systems might change the intervention—its safety, effectiveness, efficiency, and acceptability. This requires more collaboration between disciplines such as health psychology, computer science, cognitive science, health informatics, medical sociology, and public health and health promotion.

Acknowledgments

This work was funded by the Northwest Institute for BioHealth Informatics (NIBHI), UK and by EPSRC grant "e-Health+: Citizen-driven Information for Healthcare and Wellbeing" EP/G002134/1.

Conflicts of Interest

None declared.

References

1. Walshe K, Smith J. Healthcare Management. 2nd edition. London: Open University Press; 2011.
2. 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](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
3. Ryan D, Pinnock H, Tarassenko L, Lee A, Sheikh A, Price D. Can your mobile phone improve your asthma? *Thorax* 2010;65:A61. [doi: [10.1136/thx.2010.150946.36](https://doi.org/10.1136/thx.2010.150946.36)]
4. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. *Patient Educ Couns* 2009 Feb;74(2):156-173. [doi: [10.1016/j.pec.2008.08.023](https://doi.org/10.1016/j.pec.2008.08.023)] [Medline: [18947966](https://pubmed.ncbi.nlm.nih.gov/18947966/)]
5. Enwald HP, Huotari ML. Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review. *J Med Internet Res* 2010;12(2):e24 [FREE Full text] [doi: [10.2196/jmir.1409](https://doi.org/10.2196/jmir.1409)] [Medline: [20584698](https://pubmed.ncbi.nlm.nih.gov/20584698/)]
6. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. *J Med Internet Res* 2011;13(1):e2 [FREE Full text] [doi: [10.2196/jmir.1639](https://doi.org/10.2196/jmir.1639)] [Medline: [21212045](https://pubmed.ncbi.nlm.nih.gov/21212045/)]
7. Lau PW, Lau EY, Wong del P, Ransdell L. A systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents. *J Med Internet Res* 2011;13(3):e48 [FREE Full text] [doi: [10.2196/jmir.1533](https://doi.org/10.2196/jmir.1533)] [Medline: [21749967](https://pubmed.ncbi.nlm.nih.gov/21749967/)]
8. Gibbons MC, Wilson RF, Samal L, Lehmann CU, Dickersin K, Lehmann HP, et al. US Department of Health and Human Services, Agency for Healthcare Research and Quality. 2009 Oct. Impact of Consumer Health Informatics Applications. Evidence Report/Technology Assessment No 188 URL: <http://www.ahrq.gov/downloads/pub/evidence/pdf/chiapp/impactchia.pdf> [accessed 2012-03-28] [WebCite Cache ID 66V8FpQSH]

9. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. Waltham, MA: Morgan Kaufmann; 2003.
10. Lehto T, Oinas-Kukkonen H. Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. *J Med Internet Res* 2011;13(3):e46 [FREE Full text] [doi: [10.2196/jmir.1559](https://doi.org/10.2196/jmir.1559)] [Medline: [21795238](https://pubmed.ncbi.nlm.nih.gov/21795238/)]
11. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inf Syst* 2009;24:article 28.
12. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* 1997;12(1):38-48. [Medline: [10170434](https://pubmed.ncbi.nlm.nih.gov/10170434/)]
13. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
14. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977 Mar;84(2):191-215. [Medline: [847061](https://pubmed.ncbi.nlm.nih.gov/847061/)]
15. Bandura A. Social cognitive theory of self-regulation. *Organ Behav Hum Decis Process* 1991;50(2):248-287. [doi: [10.1016/0749-5978\(91\)90022-L](https://doi.org/10.1016/0749-5978(91)90022-L)]
16. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* 1991;50(2):179-211. [doi: [10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)]
17. Linn AJ, Vervloet M, van Dijk L, Smit EG, Van Weert JC. Effects of eHealth interventions on medication adherence: a systematic review of the literature. *J Med Internet Res* 2011;13(4):e103 [FREE Full text] [doi: [10.2196/jmir.1738](https://doi.org/10.2196/jmir.1738)] [Medline: [22138112](https://pubmed.ncbi.nlm.nih.gov/22138112/)]
18. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. *J Med Internet Res* 2009;11(2):e16 [FREE Full text] [doi: [10.2196/jmir.1138](https://doi.org/10.2196/jmir.1138)] [Medline: [19632970](https://pubmed.ncbi.nlm.nih.gov/19632970/)]
19. Revere D, Dunbar PJ. Review of computer-generated outpatient health behavior interventions: clinical encounters "in absentia". *J Am Med Inform Assoc* 2001 Feb;8(1):62-79 [FREE Full text] [Medline: [11141513](https://pubmed.ncbi.nlm.nih.gov/11141513/)]
20. Bickmore T, Giorgino T. Health dialog systems for patients and consumers. *J Biomed Inform* 2006 Oct;39(5):556-571. [doi: [10.1016/j.jbi.2005.12.004](https://doi.org/10.1016/j.jbi.2005.12.004)] [Medline: [16464643](https://pubmed.ncbi.nlm.nih.gov/16464643/)]
21. Corkrey R, Parkinson L. Interactive voice response: review of studies 1989-2000. *Behav Res Methods Instrum Comput* 2002 Aug;34(3):342-353. [Medline: [12395550](https://pubmed.ncbi.nlm.nih.gov/12395550/)]
22. Fernandez-Luque L, Karlsen R, Bonander J. Review of extracting information from the Social Web for health personalization. *J Med Internet Res* 2011;13(1):e15 [FREE Full text] [doi: [10.2196/jmir.1432](https://doi.org/10.2196/jmir.1432)] [Medline: [21278049](https://pubmed.ncbi.nlm.nih.gov/21278049/)]
23. 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](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
24. Howland JL, Wright TC, Boughan RA, Roberts BC. How scholarly is Google Scholar? A comparison to library databases. *Coll Res Libr* 2009;70(3):227-234.
25. Walters WH. Google Scholar coverage of a multidisciplinary field. *Inf Process Manag* 2007;43(4):1121-1132. [doi: [10.1016/j.ipm.2006.08.000](https://doi.org/10.1016/j.ipm.2006.08.000)]
26. Soureti A, Murray P, Cobain M, Chinapaw M, van Mechelen W, Hurling R. Exploratory study of web-based planning and mobile text reminders in an overweight population. *J Med Internet Res* 2011;13(4):e118 [FREE Full text] [doi: [10.2196/jmir.1773](https://doi.org/10.2196/jmir.1773)] [Medline: [22182483](https://pubmed.ncbi.nlm.nih.gov/22182483/)]
27. Ananthanarayan S, Siek KA. Health sense: a gedanken experiment on persuasive wearable technology for health awareness. New York, NY: ACM Press; 2010 Presented at: First ACM International Health Informatics Symposium: IHI '10; Nov 11-12, 2010; Arlington, VA, USA. [doi: [10.1145/1882992.1883051](https://doi.org/10.1145/1882992.1883051)]
28. Arteaga SM, Kudeki M, Woodworth A. Combating obesity trends in teenagers through persuasive mobile technology. *ACM SIGACCESS Access Comput* 2009;94:17-25. [doi: [10.1145/1595061.1595064](https://doi.org/10.1145/1595061.1595064)]
29. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. *ACM Trans Comput Hum Interact* 2005;12(2):293-327.
30. Bickmore TW, Caruso L, Clough-Gorr K, Heeren T. 'It's just like you talk to a friend' relational agents for older adults. *Interact Comput* 2005;17(6):711-735. [doi: [10.1016/j.intcom.2005.09.002](https://doi.org/10.1016/j.intcom.2005.09.002)]
31. Bickmore TW, Sidner CL. Towards plan-based health behavior change counseling systems. 2006 Presented at: AAAI Spring Symposium on Argumentation for Consumers of Healthcare; Mar 27-29, 2006; Stanford, CA, USA.
32. Bickmore TW, Mauer D, Brown T. Context awareness in a handheld exercise agent. *Pervasive Mob Comput* 2009 Jun 1;5(3):226-235. [doi: [10.1016/j.pmcj.2008.05.004](https://doi.org/10.1016/j.pmcj.2008.05.004)] [Medline: [20161031](https://pubmed.ncbi.nlm.nih.gov/20161031/)]
33. Bickmore TW, Pfeifer LM, Jack BW. Taking the time to care: empowering low health literacy patients with virtual nurse agents. In: Proceedings. 2009 Presented at: 27th International Conference on Human Factors in Computing Systems: CHI '09; Apr 4-9, 2009; Boston, MA, USA. [doi: [10.1145/1518701.1518891](https://doi.org/10.1145/1518701.1518891)]
34. Bickmore TW, Puskar K, Schlenk EA, Pfeifer LM, Sereika SM. Maintaining reality: relational agents for antipsychotic medication adherence. *Interact Comput* 2010;22(4):276-288. [doi: [10.1016/j.intcom.2010.02.001](https://doi.org/10.1016/j.intcom.2010.02.001)]
35. Bickmore TW, Schulman D, Sidner CL. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *J Biomed Inform* 2011 Apr;44(2):183-197. [doi: [10.1016/j.jbi.2010.12.006](https://doi.org/10.1016/j.jbi.2010.12.006)] [Medline: [21220044](https://pubmed.ncbi.nlm.nih.gov/21220044/)]

36. Bickmore T, Schulman D, Yin L. Maintaining Engagement in Long-term Interventions with Relational Agents. *Appl Artif Intell* 2010 Jul 1;24(6):648-666. [doi: [10.1080/08839514.2010.492259](https://doi.org/10.1080/08839514.2010.492259)] [Medline: [21318052](https://pubmed.ncbi.nlm.nih.gov/21318052/)]
37. Bieber G, Voskamp J, Urban B. Activity recognition for everyday life on mobile phones. 2009 Presented at: International Conference on Universal Access in Human Computer Interaction; Jul 19-24, 2009; San Diego, CA, USA. [doi: [10.1007/978-3-642-02710-9_32](https://doi.org/10.1007/978-3-642-02710-9_32)]
38. Buttussi F, Chittaro L. MOPET: a context-aware and user-adaptive wearable system for fitness training. *Artif Intell Med* 2008 Feb;42(2):153-163. [doi: [10.1016/j.artmed.2007.11.004](https://doi.org/10.1016/j.artmed.2007.11.004)] [Medline: [18234481](https://pubmed.ncbi.nlm.nih.gov/18234481/)]
39. Consolvo S, Klasnja P, McDonald DW, Avrahami D, Froehlich J, LeGrand L, et al. Flowers or a robot army? Encouraging awareness & activity with personal, mobile displays. 2008 Presented at: 10th International Conference on Ubiquitous Computing: UbiComp '08; Sep 21-24, 2008; Seoul, South Korea. [doi: [10.1145/1409635.1409644](https://doi.org/10.1145/1409635.1409644)]
40. de Rosi F, Novielli N, Carofiglio V, Cavalluzzi A, De Carolis B. User modeling and adaptation in health promotion dialogs with an animated character. *J Biomed Inform* 2006 Oct;39(5):514-531. [doi: [10.1016/j.jbi.2006.01.001](https://doi.org/10.1016/j.jbi.2006.01.001)] [Medline: [16524784](https://pubmed.ncbi.nlm.nih.gov/16524784/)]
41. Farzanfar R, Frishkopf S, Friedman R, Ludena K. Evaluating an automated mental health care system: making meaning of human-computer interaction. *Comput Hum Behav* 2007;23(2):1167-1182. [doi: [10.1016/j.chb.2004.11.015](https://doi.org/10.1016/j.chb.2004.11.015)]
42. Hakulinen J, Turunen M, Smith C, Cavazza M, Charlton, D. A Model for flexible interoperability between dialogue management and domain reasoning for conversational spoken dialogue systems. 2008 Presented at: Fourth International Workshop on Human-Computer Conversation; Oct 6-7, 2008; Bellagio, Italy.
43. Hartmann CW, Sciamanna CN, Blanch DC, Mui S, Lawless H, Manocchia M, et al. A website to improve asthma care by suggesting patient questions for physicians: qualitative analysis of user experiences. *J Med Internet Res* 2007;9(1):e3 [FREE Full text] [doi: [10.2196/jmir.9.1.e3](https://doi.org/10.2196/jmir.9.1.e3)] [Medline: [17478412](https://pubmed.ncbi.nlm.nih.gov/17478412/)]
44. Hayes TL, Cobbinah K, Dishongh T, Kaye JA, Kimel J, Labhard M, et al. A study of medication-taking and unobtrusive, intelligent reminding. *Telemed J E Health* 2009 Oct;15(8):770-776. [doi: [10.1089/tmj.2009.0033](https://doi.org/10.1089/tmj.2009.0033)] [Medline: [19780692](https://pubmed.ncbi.nlm.nih.gov/19780692/)]
45. Jin SA. The effects of incorporating a virtual agent in a computer-aided test designed for stress management education: the mediating role of enjoyment. *Comput Hum Behav* 2010;26(3):443-451. [doi: [10.1016/j.chb.2009.12.003](https://doi.org/10.1016/j.chb.2009.12.003)]
46. Kaipainen K, Honka A, Saranummi N. Personalized behavior change support for disease prevention. *Conf Proc IEEE Eng Med Biol Soc* 2011;2011:880-883. [doi: [10.1109/IEMBS.2011.6090196](https://doi.org/10.1109/IEMBS.2011.6090196)] [Medline: [22254451](https://pubmed.ncbi.nlm.nih.gov/22254451/)]
47. Klein M, Mogles N, van Wissen A. Why won't you do what's good for you? Using intelligent support for behavior change. 2011 Presented at: Human Behavior Understanding - Second International Workshop; Nov 16, 2011; Amsterdam, the Netherlands. [doi: [10.1007/978-3-642-25446-8_12](https://doi.org/10.1007/978-3-642-25446-8_12)]
48. Konovalov S, Scotch M, Post L, Brandt C. Biomedical informatics techniques for processing and analyzing web blogs of military service members. *J Med Internet Res* 2010;12(4):e45 [FREE Full text] [doi: [10.2196/jmir.1538](https://doi.org/10.2196/jmir.1538)] [Medline: [20923755](https://pubmed.ncbi.nlm.nih.gov/20923755/)]
49. Lee D, Helal S, Johnson B. An action-based behavior model for persuasive telehealth. In: Lee Y, Bien ZZ, Mokhtari M, Kim JT, Park M, Lee H, et al, editors. *Aging Friendly Technology for Health and Independence*, Lecture Notes in Computer Science. Berlin: Springer; 2010:121-129.
50. Levin E, Levin A. Evaluation of spoken dialogue technology for real-time health data collection. *J Med Internet Res* 2006;8(4):e30 [FREE Full text] [doi: [10.2196/jmir.8.4.e30](https://doi.org/10.2196/jmir.8.4.e30)] [Medline: [17213048](https://pubmed.ncbi.nlm.nih.gov/17213048/)]
51. Lisetti CL, Wagner E. Mental health promotion with animated characters: exploring issues and potential. 2008 Presented at: AAAI 2008 Spring Symposium on Emotion, Personality and Human Behavior; Mar 26-28, 2008; Palo Alto, CA, USA.
52. Looije R, Neerinx MA, Cnossen F. Persuasive robotic assistant for health self-management of older adults: design and evaluation of social behaviors. *Int J Hum Comput Stud* 2010;68(6):386-397. [doi: [10.1016/j.ijhcs.2009.08.007](https://doi.org/10.1016/j.ijhcs.2009.08.007)]
53. Maier E, Reimer U, Schar SG, Zimmermann PG. SEMPER: a Web-based support system for patient self-management. 2010 Presented at: 23rd Bled eConference: eTrust: Implications for the Individual, Enterprises and Society; Jun 20-23, 2010; Bled, Slovenia URL: [http://ecenter.fov.uni-mb.si/proceedings.nsf/Proceedings/16DB5E433CFFAC33C1257757003AEB6E/\\$File/15_Maier.pdf](http://ecenter.fov.uni-mb.si/proceedings.nsf/Proceedings/16DB5E433CFFAC33C1257757003AEB6E/$File/15_Maier.pdf)
54. Mazzotta I, de Rosi F, Carofiglio V. PORTIA: a user-adapted persuasion system in the healthy eating domain. *IEEE Intell Syst* 2007;22(6):42-51. [doi: [10.1109/MIS.2007.115](https://doi.org/10.1109/MIS.2007.115)]
55. Munguia Tapia E. Using Machine Learning for Real-Time Activity Recognition and Estimation of Energy Expenditure [PhD thesis]. Cambridge, MA: Massachusetts Institute of Technology; 2008. URL: <http://dspace.mit.edu/handle/1721.1/44913> [accessed 2011-07-19] [WebCite Cache ID 60IE0xNWM]
56. Nguyen H, Masthoff J. Designing persuasive dialogue systems: using argumentation with care. 2008 Presented at: Third International Conference on Persuasive Technology; Jun 4-6, 2008; Oulu, Finland. [doi: [10.1007/978-3-540-68504-3_18](https://doi.org/10.1007/978-3-540-68504-3_18)]
57. Oddsson LI, Radomski MV, White M, Nilsson D. A robotic home telehealth platform system for treatment adherence, social assistance and companionship - an overview. *Conf Proc IEEE Eng Med Biol Soc* 2009;2009:6437-6440. [doi: [10.1109/IEMBS.2009.5333744](https://doi.org/10.1109/IEMBS.2009.5333744)] [Medline: [19964427](https://pubmed.ncbi.nlm.nih.gov/19964427/)]
58. Op den Akker H, Moualed LS, Jones VM, Hermens HJ. A self-learning personalized feedback agent for motivating physical activity. 2011 Presented at: 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies: ISABEL '11; Oct 26-29, 2011; Barcelona, Spain. [doi: [10.1145/2093698.2093845](https://doi.org/10.1145/2093698.2093845)]

59. Rojas-Barahona LM, Giorgino T. Adaptable dialog architecture and runtime engine (AdaRTE): a framework for rapid prototyping of health dialog systems. *Int J Med Inform* 2009 Apr;78 Suppl 1:S56-S68. [doi: [10.1016/j.ijmedinf.2008.07.017](https://doi.org/10.1016/j.ijmedinf.2008.07.017)] [Medline: [18799352](https://pubmed.ncbi.nlm.nih.gov/18799352/)]
60. Smith C, Cavazza M, Charlton D, Zhang L, Turunen M, Hakulinen J. Integrating planning and dialogue in a lifestyle agent. In: *Intelligent Virtual Agents*. Berlin: Springer; 2008:146-153.
61. Sorbi MJ, Mak SB, Houtveen JH, Kleiboer AM, van Doornen LJ. Mobile Web-based monitoring and coaching: feasibility in chronic migraine. *J Med Internet Res* 2007;9(5):e38 [FREE Full text] [doi: [10.2196/jmir.9.5.e38](https://doi.org/10.2196/jmir.9.5.e38)] [Medline: [18166526](https://pubmed.ncbi.nlm.nih.gov/18166526/)]
62. Spring B, Schneider K, McFadden HG, Vaughn J, Kozak AT, Smith M, et al. Make Better Choices (MBC): study design of a randomized controlled trial testing optimal technology-supported change in multiple diet and physical activity risk behaviors. *BMC Public Health* 2010;10:586 [FREE Full text] [doi: [10.1186/1471-2458-10-586](https://doi.org/10.1186/1471-2458-10-586)] [Medline: [20920275](https://pubmed.ncbi.nlm.nih.gov/20920275/)]
63. Tiwari P, Warren J, Day K. Empowering older patients to engage in self care: designing an interactive robotic device. *AMIA Annu Symp Proc* 2011;2011:1402-1411. [Medline: [22195203](https://pubmed.ncbi.nlm.nih.gov/22195203/)]
64. Turunen M, Hakulinen J, Stahl O, Gambäck B, Hansen P, Rodriguez Gancedo M, et al. Multimodal and mobile conversational health and fitness companions. *Comput Speech Lang* 2011;25(2):192-209. [doi: [10.1016/j.csl.2010.04.004](https://doi.org/10.1016/j.csl.2010.04.004)]
65. Uribe JA, Duitama JF, Gomez NG. Personalized message emission in a mobile application for supporting therapeutic adherence. 2011 Presented at: 13th IEEE International Conference on e-Health Networking, Applications and Services (Healthcom); Jun 13-15, 2011; Columbia, MO, USA. [doi: [10.1109/HEALTH.2011.6026734](https://doi.org/10.1109/HEALTH.2011.6026734)]
66. van der Putten A, Kramer NC, Eimler SC. Living with a robot companion: empirical study on the interaction with an artificial health advisor. 2011 Presented at: 13th International Conference on Multimodal Interfaces: ICMI '11; Nov 14-18, 2011; Alicante, Spain. [doi: [10.1145/2070481.2070544](https://doi.org/10.1145/2070481.2070544)]
67. Watson A, Bickmore T, Cange A, Kulshreshtha A, Kvedar J. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. *J Med Internet Res* 2012;14(1):e1 [FREE Full text] [doi: [10.2196/jmir.1629](https://doi.org/10.2196/jmir.1629)] [Medline: [22281837](https://pubmed.ncbi.nlm.nih.gov/22281837/)]
68. Russell S, Norvig P. *Artificial Intelligence: A Modern Approach*. 3rd edition. Upper Saddle River, NJ: Prentice Hall; 2010.
69. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008;4:1-32. [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
70. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol* 2010 Feb;15(Pt 1):1-39. [doi: [10.1348/135910709X466063](https://doi.org/10.1348/135910709X466063)] [Medline: [19646331](https://pubmed.ncbi.nlm.nih.gov/19646331/)]
71. Miller WR, Rollnick S. *Motivational Interviewing: Preparing People to Change Addictive Behavior*. New York, NY: Guilford Press; 1991.
72. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res* 2010;12(2):e10 [FREE Full text] [doi: [10.2196/jmir.1371](https://doi.org/10.2196/jmir.1371)] [Medline: [20439251](https://pubmed.ncbi.nlm.nih.gov/20439251/)]
73. 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](https://doi.org/10.2196/jmir.1838)] [Medline: [21840837](https://pubmed.ncbi.nlm.nih.gov/21840837/)]
74. van der Krieke L, Emerencia AC, Aiello M, Sytma S. Usability evaluation of a web-based support system for people with a schizophrenia diagnosis. *J Med Internet Res* 2012;14(1):e24 [FREE Full text] [doi: [10.2196/jmir.1921](https://doi.org/10.2196/jmir.1921)] [Medline: [22311883](https://pubmed.ncbi.nlm.nih.gov/22311883/)]
75. Desmarais MC, Baker RSJ. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Model User-adapt Interact* 2011;22(1-2):9-38. [doi: [10.1007/s11257-011-9106-8](https://doi.org/10.1007/s11257-011-9106-8)]
76. Bull S, Kay J. Student models that invite the learner in: the SMILI open learner modelling framework. *Int J Artif Intellig Educ* 2007;17(2):89-120 [FREE Full text]
77. Dimitrova V. STyLE-OLM: Interactive open learner modelling. *Int J Artif Intell Educ* 2003;13(1):35-78.
78. Mitrovic A, Martin B. Evaluating the effects of open student models on learning. In: *Adaptive Hypermedia and Adaptive Web-Based Systems*. Berlin: Springer; 2006:296-305.
79. Duan D, Mitrovic A, Churcher N. Evaluating the effectiveness of multiple open student models in EER-Tutor. 2010 Presented at: 18th International Conference on Computers in Education: ICCE 2010; Nov 29-Dec 3, 2010; Putrajaya, Malaysia.
80. Du Boulay B, Avramides K, Luckin R, Martinez-Miron E, Mendez GR, Carr A. Towards systems that care: a conceptual framework based on motivation, metacognition and affect. *Int J Artif Intell Educ* 2010;20(3):197-229 [FREE Full text]
81. Mabbot A, Bull S. Student preferences for editing, persuading, and negotiating the open learner model. 2006 Presented at: 8th International Conference on Intelligent Tutoring Systems; June 36-30, 2006; Jhongli, Taiwan. [doi: [10.1007/11774303_48](https://doi.org/10.1007/11774303_48)]
82. Hughes S, Cohen D. Can online consumers contribute to drug knowledge? A mixed-methods comparison of consumer-generated and professionally controlled psychotropic medication information on the internet. *J Med Internet Res* 2011;13(3):e53 [FREE Full text] [doi: [10.2196/jmir.1716](https://doi.org/10.2196/jmir.1716)] [Medline: [21807607](https://pubmed.ncbi.nlm.nih.gov/21807607/)]
83. Wicks P, Massagli M, Kulkarni A, Dastani H. Use of an online community to develop patient-reported outcome instruments: the Multiple Sclerosis Treatment Adherence Questionnaire (MS-TAQ). *J Med Internet Res* 2011;13(1):e12 [FREE Full text] [doi: [10.2196/jmir.1687](https://doi.org/10.2196/jmir.1687)] [Medline: [21266318](https://pubmed.ncbi.nlm.nih.gov/21266318/)]

84. Gupta S, Wan FT, Newton D, Bhattacharyya OK, Chignell MH, Straus SE. WikiBuild: a new online collaboration process for multistakeholder tool development and consensus building. *J Med Internet Res* 2011;13(4):e108 [FREE Full text] [doi: [10.2196/jmir.1833](https://doi.org/10.2196/jmir.1833)] [Medline: [22155694](https://pubmed.ncbi.nlm.nih.gov/22155694/)]
85. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J Med Internet Res* 2012;14(2):e46 [FREE Full text] [doi: [10.2196/jmir.1988](https://doi.org/10.2196/jmir.1988)] [Medline: [22397809](https://pubmed.ncbi.nlm.nih.gov/22397809/)]
86. Hartzler A, Pratt W. Managing the personal side of health: how patient expertise differs from the expertise of clinicians. *J Med Internet Res* 2011;13(3):e62 [FREE Full text] [doi: [10.2196/jmir.1728](https://doi.org/10.2196/jmir.1728)] [Medline: [21846635](https://pubmed.ncbi.nlm.nih.gov/21846635/)]
87. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. *Health Educ Res* 2008 Jun;23(3):454-466 [FREE Full text] [doi: [10.1093/her/cyn004](https://doi.org/10.1093/her/cyn004)] [Medline: [18349033](https://pubmed.ncbi.nlm.nih.gov/18349033/)]
88. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007 Jul;133(4):673-693. [doi: [10.1037/0033-2909.133.4.673](https://doi.org/10.1037/0033-2909.133.4.673)] [Medline: [17592961](https://pubmed.ncbi.nlm.nih.gov/17592961/)]

Abbreviations

ECA: embodied conversational agent

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 05.09.11; peer-reviewed by T Bickmore; comments to author 25.09.11; revised version received 31.03.12; accepted 15.05.12; published 14.06.12.

Please cite as:

Kennedy CM, Powell J, Payne TH, Ainsworth J, Boyd A, Buchan I

Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review

J Med Internet Res 2012;14(3):e80

URL: <http://www.jmir.org/2012/3/e80/>

doi: [10.2196/jmir.1893](https://doi.org/10.2196/jmir.1893)

PMID: [22698679](https://pubmed.ncbi.nlm.nih.gov/22698679/)

©Catriona M Kennedy, John Powell, Thomas H Payne, John Ainsworth, Alan Boyd, Iain Buchan. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 14.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Editorial

Beyond Efficacy: The Depth and Diversity of Current Internet Interventions

Helen Christensen¹, PhD; Alison L Cascar², PhD; Gerhard Andersson³, PhD; Frances P Thorndike⁴, PhD; Robert J Tait², PhD

¹Black Dog Institute, University of New South Wales, Sydney, NSW, Australia

²Centre for Mental Health Research, The Australian National University, Canberra, Australia

³Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

⁴Department of Psychiatry and Neurobehavioral Sciences, University of Virginia Health System, Charlottesville, VA, United States

Corresponding Author:

Helen Christensen, PhD

Black Dog Institute

University of New South Wales

Hospital Road

Randwick

Sydney, NSW, 2031

Australia

Phone: 61 2 9382 9288

Fax: 61 2 9382 8208

Email: h.christensen@unsw.edu.au

(*J Med Internet Res* 2012;14(3):e92) doi:[10.2196/jmir.2206](https://doi.org/10.2196/jmir.2206)

This issue of the *Journal of Medical Internet Research* (JMIR) brings together nine papers selected from presentations at the 5th International Conference of the *International Society for Research on Internet Interventions (ISRII)* held in Sydney, Australia in 2011. The papers highlight the depth and diversity of our field and reflect some of the current major research themes, including issues of stigma, new modes of intervention delivery, tailoring, the role of support, mechanisms of change, target audience considerations, new theoretical approaches, and further grounding in models and frameworks. The field is pushing beyond efficacy trials to better understand when, how, and why Internet interventions work. As was true of many of the presentations at ISRII, this set of papers makes a substantive contribution to our field's understanding of how to reduce mental and behavioral health problems (for example, depression, anxiety, smoking), improve mental health literacy and knowledge, and encourage help-seeking. In the remaining paragraphs, we summarize several key points on how each paper contributes to the broader research picture in these areas, suggesting future paths of research.

Stigmatizing attitudes and beliefs towards mental health disorders is a significant problem in the community, with mental health stigma associated with poor help-seeking behavior and treatment adherence. A number of effective programs have been developed to address mental health stigma in the community. These have included programs aimed at challenging stereotypes and increasing knowledge through education and/or personal contact with a person with a mental health problem [1, 2]. The paper by Gulliver et al. [3] presents the efficacy of two brief

online interventions to promote help-seeking behavior in elite athletes, with the study suggesting that an online mental health literacy intervention can reduce depression and anxiety stigma. Online interventions such as this one have the potential to reach a broad audience and may be a cost-effective way of reducing stigma in the community. The paper by Farrer and colleagues [4] further supports the effectiveness of online interventions in reducing depression stigma, finding significantly lower levels of stigma amongst participants completing an online cognitive behavior therapy (CBT) program for depression.

Kauer et al. [5] continued the field's investigation of using technology to treat depression, but examined the use of mobile phone self-monitoring in a group of adolescents with mild depressive symptoms. The authors found that self-monitoring increased emotional self-awareness which in turn mediated decreases in depressive symptoms. Monitoring symptoms and activities by means of mobile phone technology is likely to become even more common in research given increased use of smart phones in daily life. Smart phones are also extending the reach of these interventions to patient groups with historically limited Internet access.

Tailored Internet interventions are also being more readily developed and studied. Rather than offering the same intervention to all users, harnessing the full capabilities of the technology allows these interventions to be tailored to specific user characteristics (for example, symptom level, age). Silfvernagel et al. [6] report data from a controlled trial in which individually tailored, guided cognitive-behavior therapy was tested for persons with panic symptoms. Findings suggest that

tailoring the Internet intervention to this patient population was feasible and efficacious, adding to the previous promising findings with large effect sizes [7, 8].

The role of support in Internet interventions has been discussed over the years, with considerable debate about whether support is necessary to achieve optimal outcomes, and, if so, at what cost [7, 9-14]. Support can be provided via multiple formats (for example, personalized emails, texts, phone calls) and can occur at various points in the intervention. Three papers in this special edition address the issue of support by exploring whether clinician delivered, clinician guided, guided or self-guided interventions are equally effective, as well as for whom and when (see [4, 6, 15]). The trial by Farrer et al. was conducted in a national helpline setting and compared outcomes among participants who received: 1) web-based CBT for depression; 2) web-based CBT plus telephone tracking; 3) weekly telephone tracking only; or 3) neither web CBT nor telephone tracking. In this study, telephone support during the actual treatment did not affect outcome, indicating that the temporal aspects of support and contact need to be further investigated [4]. In the Silfvernagel et al. paper [6] mentioned above, the tailored intervention for panic used online therapist guidance to support users. Overall, the role of support and better definitions of precisely what support entails, warrants further evaluation.

Anderson and colleagues [15] examined a therapist guided Internet intervention in which adolescents and children with anxiety disorders, as well as their parents, were included. Borrowing a concept from traditional face-to-face psychotherapy, the authors investigated the role of working alliance, a term to describe the collaboration between a patient and a therapist. In line with previous studies on adult patients [16, 17], high ratings of working alliance were found, even equivalent to what was observed in clinic-based treatment. Alliance also predicted adherence and outcome for the adolescents but not for the younger children. The findings from this study highlight the importance of investigating differences and similarities between traditional and Internet interventions, as well as the importance of examining mechanisms of change within Internet interventions.

The use of Internet interventions as a means of addressing problematic behaviors, including substance use, is now well established. A review by Riper and colleagues [18] found that even a single session achieves effects of a similar magnitude to that found for in-person brief alcohol interventions [19] with multi-session interventions producing significantly larger effects. However, much of the research in this field has utilized university samples that can be expected to have higher levels of computer literacy than the general community [20], leaving the effectiveness of Internet interventions in larger populations open to question. The paper by Muñoz et al. addresses this issue, at least in the context of the cessation of smoking [21]. Having demonstrated the efficacy of the combined Spanish / English language San Francisco Stop Smoking Internet site, the resource was opened to the general public (aged 18 or over). With participants from more than 150 counties, the observed quit rate

at 12 months (45%) was higher than the baseline trial and incurred minimal costs. High rates of substance use frequently co-occur with other mental health problems [22]. Although, there have been interventions developed specifically addressing those with comorbid alcohol and depression [23], the paper by Farrer et al. [4] shows significant short-term reductions in hazardous alcohol use as a secondary effect of an intervention for depression.

An alternative approach to smoking cessation is through the use of online social networks. Popular components of Web-Assisted Tobacco Interventions (WATIs) are social support networks or discussion boards [24]. The paper by van Mierlo and colleagues [24] provides an analysis of the categories of users who post to discussion boards on WATIs. Similar typographies of users were found from a publicly funded resource (*Smokers Helpline Online*) and from a social enterprise site (*StopSmokingCenter.net*). Of particular interest was the finding that a sub-group (termed “superusers”) who represented less than 1% of registrants accounted for 35-45% of all posts. An earlier review noted the paucity of evidence for the effectiveness of Internet Support Groups [25]. The findings by the van Mierlo team highlight a particular group for specific research attention in the future, given their likely impact on the tone and outcomes of discussion boards.

Another theme within Internet intervention research is consideration of the target audience; in this special issue, some papers evaluate treatment interventions targeted at symptomatic individuals (see [5, 6, 15]) and others examine the value of public health interventions (see [21, 26]), including in the novel area of online social networks (see [24]). Within Internet-delivered interventions, the major approach remains cognitive behavioral therapy (see [5, 6, 15]), but research is also being conducted with different theoretical approaches, including positive psychology (see [26]). The field is also examining its grounding within various theoretical frameworks, and the paper written by Hilgart and colleagues proposes using a proven methodology (instructional design models) to guide the planning, development, and implementation of Internet interventions [27].

In conclusion, this theme issue demonstrates the breadth of research being conducted within the *International Society for Research on Internet Interventions*. It shows how rapidly the field is moving forward, pushing beyond initial feasibility and efficacy trials to more nuanced questions, such as the testing of different levels of support, various technologies, potential mechanisms of change, and implementation on a public health scale. The field is well grounded in theoretical models of behavior change, but is calling for more theory-based intervention planning, design, and development. The rich array of research questions and investigations in this special edition shows the field is advancing to not only raising questions for specific populations or specific intervention types, but instead raising research questions for the field as a whole. The papers clearly suggest new lines of research for the future and show the promise of this emerging field.

Conflicts of Interest

None declared.

References

1. Pinto-Foltz MD, Logsdon MC. Reducing stigma related to mental disorders: initiatives, interventions, and recommendations for nursing. *Arch Psychiatr Nurs* 2009 Feb;23(1):32-40. [doi: [10.1016/j.apnu.2008.02.010](https://doi.org/10.1016/j.apnu.2008.02.010)] [Medline: [19216986](https://pubmed.ncbi.nlm.nih.gov/19216986/)]
2. Schachter HM, Girardi A, Ly M, Lacroix D, Lumb AB, van Berkum J, et al. Effects of school-based interventions on mental health stigmatization: a systematic review. *Child Adolesc Psychiatry Ment Health* 2008;2(1):18 [FREE Full text] [doi: [10.1186/1753-2000-2-18](https://doi.org/10.1186/1753-2000-2-18)] [Medline: [18644150](https://pubmed.ncbi.nlm.nih.gov/18644150/)]
3. Gulliver A, Griffiths KM, Christensen H, Mackinnon A, Cascar AL, Parsons A, et al. Internet-Based Interventions to Promote Mental Health Help-Seeking in Elite Athletes: An Exploratory Randomized Controlled Trial. *J Med Internet Res* 2012;14(3):e69 [FREE Full text] [doi: [10.2196/jmir.1864](https://doi.org/10.2196/jmir.1864)]
4. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Web-Based Cognitive Behavior Therapy for Depression With and Without Telephone Tracking in a National Helpline: Secondary Outcomes From a Randomized Controlled Trial. *J Med Internet Res* 2012;14(3):e68 [FREE Full text] [doi: [10.2196/jmir.1859](https://doi.org/10.2196/jmir.1859)]
5. Kauer SD, Reid SC, Crooke AHD, Khor A, Hearps SJC, Jorm AF, et al. Self-monitoring Using Mobile Phones in the Early Stages of Adolescent Depression: Randomized Controlled Trial. *J Med Internet Res* 2012;14(3):e67 [FREE Full text] [doi: [10.2196/jmir.1858](https://doi.org/10.2196/jmir.1858)]
6. Silfvernagel K, Carlbring P, Kåbo J, Edström S, Eriksson J, Månson L, et al. Individually Tailored Internet-Based Treatment for Young Adults and Adults With Panic Attacks: Randomized Controlled Trial. *J Med Internet Res* 2012;14(3):e65 [FREE Full text] [doi: [10.2196/jmir.1853](https://doi.org/10.2196/jmir.1853)]
7. Carlbring P, Maurin L, Törngren C, Linna E, Eriksson T, Sparthar E, et al. Individually-tailored, Internet-based treatment for anxiety disorders: A randomized controlled trial. *Behav Res Ther* 2011 Jan;49(1):18-24. [doi: [10.1016/j.brat.2010.10.002](https://doi.org/10.1016/j.brat.2010.10.002)] [Medline: [21047620](https://pubmed.ncbi.nlm.nih.gov/21047620/)]
8. Johansson R, Sjöberg E, Sjögren M, Johnsson E, Carlbring P, Andersson T, et al. Tailored vs. Standardized Internet-Based Cognitive Behavior Therapy for Depression and Comorbid Symptoms: A Randomized Controlled Trial. *PLoS One* 2012;7(5):e36905 [FREE Full text] [doi: [10.1371/journal.pone.0036905](https://doi.org/10.1371/journal.pone.0036905)] [Medline: [22615841](https://pubmed.ncbi.nlm.nih.gov/22615841/)]
9. Andersson G, Lundström P, Ström L. Internet-based treatment of headache: does telephone contact add anything? *Headache* 2003 Apr;43(4):353-361. [Medline: [12656706](https://pubmed.ncbi.nlm.nih.gov/12656706/)]
10. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
11. Buhrman M, Fälden S, Ström L, Andersson G. Controlled trial of Internet-based treatment with telephone support for chronic back pain. *Pain* 2004 Oct;111(3):368-377. [doi: [10.1016/j.pain.2004.07.021](https://doi.org/10.1016/j.pain.2004.07.021)] [Medline: [15363881](https://pubmed.ncbi.nlm.nih.gov/15363881/)]
12. Christensen H, Griffiths KM, Mackinnon AJ, Brittcliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. *Psychol Med* 2006 Dec;36(12):1737-1746. [doi: [10.1017/S0033291706008695](https://doi.org/10.1017/S0033291706008695)] [Medline: [16938144](https://pubmed.ncbi.nlm.nih.gov/16938144/)]
13. Kenardy J, McCafferty K, Rosa V. Internet-delivered indicated prevention for anxiety disorders: a randomized controlled trial. *Behavioural and Cognitive Psychotherapy* 2003;31(3):279-283.
14. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. *J Med Internet Res* 2010;12(5):e60 [FREE Full text] [doi: [10.2196/jmir.1463](https://doi.org/10.2196/jmir.1463)] [Medline: [21169162](https://pubmed.ncbi.nlm.nih.gov/21169162/)]
15. Anderson REE, Spence SH, Donovan CL, March S, Prosser S, Kenardy J. Working Alliance in Online Cognitive Behavior Therapy for Anxiety Disorders in Youth: Comparison With Clinic Delivery and its Role in Predicting Outcome. *J Med Internet Res* 2012;14(3):e88 [FREE Full text] [doi: [10.2196/jmir.1848](https://doi.org/10.2196/jmir.1848)]
16. Klein B, Mitchell J, Gilson K, Shandley K, Austin D, Kiropoulos L, et al. A therapist-assisted Internet-based CBT intervention for posttraumatic stress disorder: preliminary results. *Cogn Behav Ther* 2009 Jun;38(2):121-131. [doi: [10.1080/16506070902803483](https://doi.org/10.1080/16506070902803483)] [Medline: [20183691](https://pubmed.ncbi.nlm.nih.gov/20183691/)]
17. Knaevelsrud C, Maercker A. Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: a randomized controlled clinical trial. *BMC Psychiatry* 2007;7:13 [FREE Full text] [doi: [10.1186/1471-244X-7-13](https://doi.org/10.1186/1471-244X-7-13)] [Medline: [17442125](https://pubmed.ncbi.nlm.nih.gov/17442125/)]
18. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. *J Med Internet Res* 2011;13(2):e42 [FREE Full text] [doi: [10.2196/jmir.1691](https://doi.org/10.2196/jmir.1691)] [Medline: [21719411](https://pubmed.ncbi.nlm.nih.gov/21719411/)]
19. Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction* 2002 Mar;97(3):279-292. [Medline: [11964101](https://pubmed.ncbi.nlm.nih.gov/11964101/)]
20. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. *J Med Internet Res* 2010;12(5):e62 [FREE Full text] [doi: [10.2196/jmir.1479](https://doi.org/10.2196/jmir.1479)] [Medline: [21169175](https://pubmed.ncbi.nlm.nih.gov/21169175/)]

21. Muñoz RF, Aguilera A, Schueller SM, Leykin Y, Pérez-Stable EJ. From Online Randomized Controlled Trials to Participant Preference Studies: Morphing the San Francisco Stop Smoking Site into a Worldwide Smoking Cessation Resource. *J Med Internet Res* 2012;14(3):e64 [FREE Full text] [doi: [10.2196/jmir.1852](https://doi.org/10.2196/jmir.1852)]
22. Kessler RC, Crum RM, Warner LA, Nelson CB, Schulenberg J, Anthony JC. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry* 1997 Apr;54(4):313-321. [Medline: [9107147](https://pubmed.ncbi.nlm.nih.gov/9107147/)]
23. Kay-Lambkin FJ, Baker AL, Lewin TJ, Carr VJ. Computer-based psychological treatment for comorbid depression and problematic alcohol and/or cannabis use: a randomized controlled trial of clinical efficacy. *Addiction* 2009 Mar;104(3):378-388. [doi: [10.1111/j.1360-0443.2008.02444.x](https://doi.org/10.1111/j.1360-0443.2008.02444.x)] [Medline: [19207345](https://pubmed.ncbi.nlm.nih.gov/19207345/)]
24. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in Social Networks for Smoking Cessation: Analysis of Demographic Characteristics and Posting Behavior From the Canadian Cancer Society's Smokers' Helpline Online and StopSmokingCenter.net. *J Med Internet Res* 2012;14(3):e66 [FREE Full text] [doi: [10.2196/jmir.1854](https://doi.org/10.2196/jmir.1854)]
25. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? *J Med Internet Res* 2009;11(3):e40 [FREE Full text] [doi: [10.2196/jmir.1270](https://doi.org/10.2196/jmir.1270)] [Medline: [19793719](https://pubmed.ncbi.nlm.nih.gov/19793719/)]
26. Schueller SM, Parks AC. Disseminating Self-Help: Positive Psychology Exercises in an Online Trial. *J Med Internet Res* 2012;14(3):e63 [FREE Full text] [doi: [10.2196/jmir.1850](https://doi.org/10.2196/jmir.1850)]
27. Hilgart MM, Ritterband LM, Thorndike FP, Kinzie MB. Using Instructional Design Process to Improve Design and Development of Internet Interventions. *J Med Internet Res* 2012;14(3):e89 [FREE Full text] [doi: [10.2196/jmir.1890](https://doi.org/10.2196/jmir.1890)]

Edited by G Eysenbach; submitted 04.06.12; this is a non-peer-reviewed article; accepted 11.06.12; published 29.06.12.

Please cite as:

Christensen H, Calear AL, Andersson G, Thorndike FP, Tait RJ
Beyond Efficacy: The Depth and Diversity of Current Internet Interventions
J Med Internet Res 2012;14(3):e92
URL: <http://www.jmir.org/2012/3/e92/>
doi: [10.2196/jmir.2206](https://doi.org/10.2196/jmir.2206)
PMID: [22743601](https://pubmed.ncbi.nlm.nih.gov/22743601/)

©Helen Christensen, Alison L Calear, Gerhard Andersson, Frances P Thorndike, Robert J Tait. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of a Web-Based Multiple Tailored Smoking Cessation Program: A Randomized Controlled Trial Among Dutch Adult Smokers

Eline Suzanne Smit^{1,2}, MSc; Hein de Vries^{1,2}, PhD; Ciska Hoving^{1,2}, PhD

¹Department of Health Promotion, Maastricht University, Maastricht, Netherlands

²CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, Netherlands

Corresponding Author:

Eline Suzanne Smit, MSc

Department of Health Promotion

Maastricht University

PO Box 616

Maastricht, 6200 MD

Netherlands

Phone: 31 433882397

Fax: 31 433671032

Email: es.smit@maastrichtuniversity.nl

Abstract

Background: Distributing a multiple computer-tailored smoking cessation intervention through the Internet has several advantages for both provider and receiver. Most important, a large audience of smokers can be reached while a highly individualized and personal form of feedback can be maintained. However, such a smoking cessation program has yet to be developed and implemented in the Netherlands.

Objective: To investigate the effects of a Web-based multiple computer-tailored smoking cessation program on smoking cessation outcomes in a sample of Dutch adult smokers.

Methods: Smokers were recruited from December 2009 to June 2010 by advertising our study in the mass media and on the Internet. Those interested and motivated to quit smoking within 6 months ($N = 1123$) were randomly assigned to either the experimental ($n = 552$) or control group ($n = 571$). Respondents in the experimental group received the fully automated Web-based smoking cessation program, while respondents in the control group received no intervention. After 6 weeks and after 6 months, we assessed the effect of the intervention on self-reported 24-hour point prevalence abstinence, 7-day point prevalence abstinence, and prolonged abstinence using logistic regression analyses.

Results: Of the 1123 respondents, 449 (40.0%) completed the 6-week follow-up questionnaire and 291 (25.9%) completed the 6-month follow-up questionnaire. We used a negative scenario to replace missing values. That is, we considered respondents lost to follow-up to still be smoking. The computer-tailored program appeared to have significantly increased 24-hour point prevalence abstinence (odds ratio [OR] 1.85, 95% confidence interval [CI] 1.30–2.65), 7-day point prevalence abstinence (OR 2.17, 95% CI 1.44–3.27), and prolonged abstinence (OR 1.99, 95% CI 1.28–3.09) rates reported after 6 weeks. After 6 months, however, no intervention effects could be identified. Results from complete-case analyses were similar.

Conclusions: The results presented suggest that the Web-based computer-tailored smoking cessation program had a significant effect on abstinence reported after a 6-week period. At the 6-month follow-up, however, no intervention effects could be identified. This might be explained by the replacement of missing values on the primary outcome measures due to attrition using a negative scenario. While results were similar when using a less conservative scenario (ie, complete-case analyses), the results should still be interpreted with caution. Further research should aim at identifying strategies that will prevent high attrition in the first place and, subsequently, to identify the best strategies for dealing with missing data when studies have high attrition rates.

Trial Registration: Dutch Trial Register NTR1351; <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=1351> (Archived by WebCite at <http://www.webcitation.org/67egSTWrz>)

(*J Med Internet Res* 2012;14(3):e82) doi:[10.2196/jmir.1812](https://doi.org/10.2196/jmir.1812)

KEYWORDS

Smoking cessation; Web-based intervention; computer tailoring; iterative feedback; intervention; randomized controlled trial

Introduction

Worldwide, the smoking of tobacco is the most preventable cause of illness and premature death [1]. Therefore, many interventions have been developed aimed at helping smokers to quit. One strategy that has shown both short- and long-term efficacy in changing smoking behavior is computer tailoring [2-7]. The content of a computer-tailored intervention is adapted to the specific characteristics of a particular individual. This has been shown to attract and keep an individual's attention [8,9], resulting in more thorough processing of information [10]. A single tailored feedback message has already proven to be effective in promoting abstinence from smoking [11], but when tailored information is provided on multiple occasions the impact of the intervention can be increased even more [2,12]. In addition, a recent meta-analysis found that dynamically tailored interventions (ie, iterative assessments and feedback) resulted in greater changes in health behavior than statically tailored interventions [13]. As this effect could not be explained solely by the increased number of overall contacts that dynamic tailoring necessitates [13], a computer-tailoring approach may be warranted that consists not only of multiple feedback moments but also of feedback messages that are iterative: the feedback that respondents receive later on during the intervention should not only concern the respondent's present state, but also refer to the changes respondents have made since their enrollment in the program.

The Internet has been discovered to be a popular gateway for delivering health behavior change interventions in general [14] and computer-tailored and smoking cessation interventions in particular [15,16]. Using the Internet to provide such programs may have several advantages for both provider and receiver: it is highly accessible [17,18], it has the potential to reach a large audience at minimal cost, and participants can take part at any time that is most convenient to them. Furthermore, the behavioral feedback given in computer-tailoring programs can still be highly individualized and personal [19]. Moreover, smokers might not succeed the first time they try to quit [20], and seeking help online relatively anonymously may prevent them from feelings of failure and embarrassment, negative feelings that have been shown to be related to a higher temptation to smoke [21] and poorer abstinence outcomes [22].

Although a key element of computer tailoring is that the intervention materials are adapted to specific respondent characteristics, some smokers might benefit more than others from particular smoking cessation interventions. For example, the level of nicotine dependence has previously been suggested to moderate the effectiveness of smoking cessation interventions [23]. Moreover, a study among nicotine patch users identified several participant characteristics moderating the effectiveness of a Web-based computer-tailored intervention [24]. It has therefore been recommended, especially with computer-tailored interventions using new media technologies such as the Internet, to investigate which participant characteristics are associated with effectiveness [4].

As Web-based multiple tailored smoking cessation feedback has not yet been offered to the Dutch general public outside scientific studies, our research group developed a Web-based multiple computer-tailored smoking cessation program and offered Dutch adult smokers the opportunity to participate in this program. The present study investigated the effectiveness of this program on smoking cessation outcomes reported after 6 weeks and 6 months. To imitate a natural situation in which smokers who do not participate in a smoking cessation program do not receive the intervention, the control group did not receive any of the intervention's components. Nevertheless, both the intervention and control group were free to use other smoking cessation aids during the study period. In addition, we investigated whether the effect of the intervention was different for specific subgroups of smokers and whether we could detect a dose-response relationship between the number of feedback messages received and abstinence at the last follow-up.

Methods**Intervention**

The Web-based multiple computer-tailored smoking cessation program was based on a previously developed effective single computer-tailored intervention [6], while the I-Change model served as the theoretical framework [25]. While filling out the online baseline questionnaire, all respondents were asked to set a quit date within the next 4 weeks. Respondents in both intervention arms were prompted by email to fill in an online follow-up questionnaire 2 days after their set quit date, and after 6 weeks and 6 months. By clicking on a link provided in this email, respondents could start filling out their next follow-up questionnaire immediately, by logging into the system. One email reminder was sent each time that, 1 week after receiving the first invitation, a respondent had still not filled out the particular questionnaire he or she was invited to complete. While respondents in the control group filled out only the questions, for those in the intervention group questions were directly succeeded by relevant online feedback in order to maintain the respondent's attention and to improve retention rates. Iterative and item-based feedback messages were tailored to several respondent characteristics [25]: gender, cognitive variables (attitude, social influence, and self-efficacy), intention to quit smoking, goal and relapse prevention strategies (action and coping plans), and smoking behavior. When the questionnaire was completed, feedback messages were combined into one personalized feedback letter. In addition to being able to read the feedback letters on the computer screen, respondents were also sent the feedback letters by email, which allowed for the letters to be printed. The 4- to 5-page feedback letters respondents received at baseline and after 6 weeks consisted of seven components: (1) introduction, including specific feedback on the respondent's smoking behavior and on his or her intention to quit smoking or to maintain nonsmoking, (2) feedback on the respondent's attitude (perceived advantages [pros] and disadvantages [cons]) toward smoking and quitting smoking, (3) feedback on perceived social influence (not) to smoke, (4)

feedback on the respondent's reported self-efficacy to refrain from smoking in specific situations, including suggestions on how to cope with these situations, (5) feedback on the extent to which respondents were planning to undertake specific actions (action plans) while preparing their quit attempt, (6) feedback on how to cope with certain difficult situations (coping plans), including the formulation of personal plans in the shape of if-then statements [26], and (7) ending. As we wanted to minimize the burden of filling out a questionnaire by smokers who had recently quit, the feedback letter that respondents received 2 days after the set quit date consisted of only 1 page,

giving feedback on smoking (cessation) behavior and relapse prevention strategies. Figure 1 shows an example of exemplary items regarding the pros of smoking cessation. Figure 2 shows an example of a tailored feedback message.

All respondents in the experimental condition received at least one tailored feedback letter (ie, at baseline). At the 6-week follow-up, respondents could have received at most two tailored feedback letters (ie, at baseline and 2 days after their set quit date), and at the 6-month follow-up, they could have received a maximum of three tailored feedback letters (ie, at baseline, 2 days after their set quit date, and at the 6-week follow-up).

Figure 1. Screenshot of items regarding the pros of smoking cessation.

Figure 2. Screenshot of personal advice regarding the pros of smoking cessation.

Recruitment and Procedure

This study was approved by the Medical Ethics Committee of Maastricht University and the University Hospital Maastricht (MEC 08-3-037; NL22692.068.08), and is registered with the Dutch Trial Register (NTR1351). A full description of the study protocol is provided elsewhere [27].

We recruited adult smokers from December 2009 to June 2010 by advertising our study in the mass media and on the Internet. We sent several press releases to regional newspapers in the Netherlands. Most of these newspapers subsequently mentioned our study on their website, included an item about the project in the print version of their newspaper, or mentioned our study

on their local radio station or television channel, or both. We also used a Dutch online social network website (Hyves) and several online smoking cessation forums to disseminate our recruitment text. In addition, we advertised our study in a free national newspaper obtainable at all Dutch train stations and several other public places throughout the Netherlands.

After 12 months, we expected a 10% point prevalence abstinence rate in the control condition. Based on results from previous projects, we expected the multiple tailoring program to lead to a 20% point prevalence abstinence rate. To be able to detect this difference significantly ($\alpha = 5\%$, $\beta = 10\%$), according to a 2-tailed Fisher exact test, 281 respondents per arm were required at the end of the trial (562 respondents in

total) [28]. Allowing for 50% attrition over the trial period, 1124 respondents needed to be included at baseline.

Interested smokers could sign up for the study on the study website (<http://www.persoonlijkstopadvies.nl>) and were eligible to participate if they were 18 years of age or older, were motivated to quit smoking within 6 months, and had access to the Internet. On the study website, participants were informed that the study was financed by the Dutch Cancer Society and conducted by researchers from Maastricht University in cooperation with the Dutch Expert Center on Tobacco Control (STIVORO). Additionally, the website included information about the objectives of the study, the randomization procedure and the incentive provided when respondents completed all questionnaires (ie, a €10 voucher). Respondents could choose their own username and password and were informed that no one but the research team was able to retrieve these passwords. As respondents had to report their email address when signing up for the study, we could easily flag respondents with multiple identities and remove them from further analyses. After providing online informed consent, participants were randomly assigned to the intervention group or the control group by a computer software randomization device, allocating approximately 50% of all respondents to each group. Blinding of respondents was not possible, as they had to take notice of whether they were receiving tailored feedback.

Measurements

All questionnaires used in the present study were previously used and tested among Dutch smoking adults and were self-administered online [6,28,29].

Baseline Measurement

We measured six demographic variables: age, gender (1 = male, 2 = female), educational level (1 = low: primary school/basic vocational school, 2 = medium: secondary vocational school/high school degree, 3 = high: higher vocational school/college degree/university degree), nationality (1 = Dutch, 2 = non-Dutch), and the occurrence of cardiovascular and respiratory diseases (1 = no, 2 = yes).

Exclusion criteria were based on current smoking behavior and motivation to quit smoking: current smoking behavior was measured by 1 item asking whether the respondent had smoked during the past 7 days (1 = no, 2 = yes). Motivation to quit smoking was measured by an adapted version of the Stage of Change algorithm [30]. We asked respondents to state within what time span they intended to quit smoking (1 = not within 6 months, 2 = within 6 months, 3 = within 1 month, 4 = I have quit, but no longer than 6 months, 5 = I have quit for longer than 6 months). Respondents who indicated that they had not smoked during the past 7 days, who were not willing to quit within 6 months, or who had quit already were excluded from further participation.

We measured overall tobacco consumption using five open-ended questions regarding the number of cigarettes, hand-rolled cigarettes, cigars, cigarillos, and pipes smoked per day. Subsequently, the answers on these five questions were converted into an overall score for tobacco consumption (expressed as number of cigarettes), whereby 1 hand-rolled

cigarette or cigarillo equaled 1 cigarette and 1 cigar equaled 4 cigarettes [31]. As no concrete guidelines were available for the number of cigarettes that would equal 1 pipe, we conservatively considered 1 pipe to equal 1 cigarette.

We measured addiction level by the abbreviated Fagerström Test for Nicotine Dependence (0 = not addicted, 10 = highly addicted) [32].

We assessed the number of past quit attempts with 1 item, asking the respondents how often they had tried to quit smoking in the past.

Follow-up Measurement

At the 6-week and 6-month follow-ups, we assessed prolonged abstinence by 1 item asking whether the respondent had refrained from smoking since the previous measurement (1 = no, 2 = yes). At the 6-week follow-up, prolonged abstinence referred to abstinence since the questionnaire that respondents received 2 days after their set quit date (ie, at least 2 weeks of abstinence). At the 6-month follow-up, this measure referred to abstinence since the 6-week follow-up (ie, 4.5 months of abstinence). In addition, at both follow-ups we assessed 24-hour and 7-day point prevalence abstinence, each by 1 item asking whether the respondent had refrained from smoking during the past 24 hours or 7 days (1 = no, 2 = yes).

Statistical Analyses

First, we conducted descriptive analyses to determine the sample's characteristics. To check for differences between the intervention and control groups, we conducted 2-sided *t* tests and chi-square tests. Additionally, to determine whether selective dropout had occurred, we compared those remaining in the study versus those lost to follow-up after 6 weeks and 6 months using 2-sided *t* tests and chi-square tests.

Second, we conducted logistic regression analyses to determine whether the intervention had an effect on the outcome measures assessed after follow-up periods of 6 weeks and 6 months. A negative scenario was used to replace missing values. That is, respondents lost to follow-up were considered to still be smoking. To test the robustness of the results, these analyses were also conducted with complete cases only.

Third, to determine whether the effect of the intervention was different for specific subgroups of smokers, we investigated whether we could identify interaction effects between the study condition and baseline demographic or behavioral measures using logistic regression analyses.

Data were analyzed using SPSS 17.0 (IBM Corporation, Somers, NY, USA). The significance level used was $P < .05$.

Results

Sample Characteristics

Figure 3 shows the flow of respondents from enrollment in the study to allocation to the experimental and control conditions, retention, and whether they were included in the analysis. Of the 1257 respondents assessed for eligibility, 33 (3%) declined to participate, 32 (3%) were nonsmokers at baseline, and 69 (6%) were not motivated to quit within 6 months. Ultimately,

1123 (89.34%) respondents were randomly assigned to either the experimental ($n = 552$) or control ($n = 571$) group and completed the baseline questionnaire. Of the 1123 respondents included, 449 (40.0%) completed the 6-week follow-up questionnaire and 291 (25.9%) completed the questionnaire at the 6-month follow-up.

Respondents included in the analyses had a mean age of 49.5 years; 535 (47.6%) were male; and 513 (45.7%) had a medium

level of education. Respondents in the experimental group significantly differed from those in the control condition in their level of education ($\chi^2_2 = 6.11$, $P = .047$). Therefore, educational level was included in subsequent analyses as a potential confounder. As no data concerning level of education were missing for any of the respondents, we included all 1123 in further analyses. Table 1 shows the baseline characteristics of the overall sample and of the experimental and control groups separately.

Table 1. Baseline sample characteristics of Dutch smoking adults ($N = 1123$) recruited from December 2009 to June 2010.

Characteristic	Overall sample ($N = 1123$)	Experimental group ($n = 552$)	Control group ($n = 571$)
Age (years), mean (SD)	49.5 (32.5)	48.4 (12.2)	48.8 (12.3)
Male, % (n)	47.6% (535)	45.8% (253)	49.4% (282)
Educational level, % (n)			
High	21.2% (238)	19.6% (108)	22.8% (130)
Medium	45.7% (513)	43.8% (242)	47.5% (271)
Low	33.1% (372)	36.6% (202)	29.8% (170)
Dutch, % (n)	97.7% (1097)	97.8% (540)	97.5% (557)
With cardiovascular diseases, % (n)	9.4% (106)	11.1% (61)	7.9% (45)
With respiratory diseases, % (n)	14.3% (161)	12.5% (69)	16.1% (92)
Number of cigarettes smoked/day, mean (SD)	20.6 (12.4)	20.8 (13.7)	20.4 (11.0)
FTND ^a score (range 1–10), mean (SD)	5.1 (2.5)	5.0 (2.5)	5.2 (2.4)
Number of previous quit attempts, mean (SD)	5.4 (17.5)	5.1 (10.1)	5.7 (22.4)

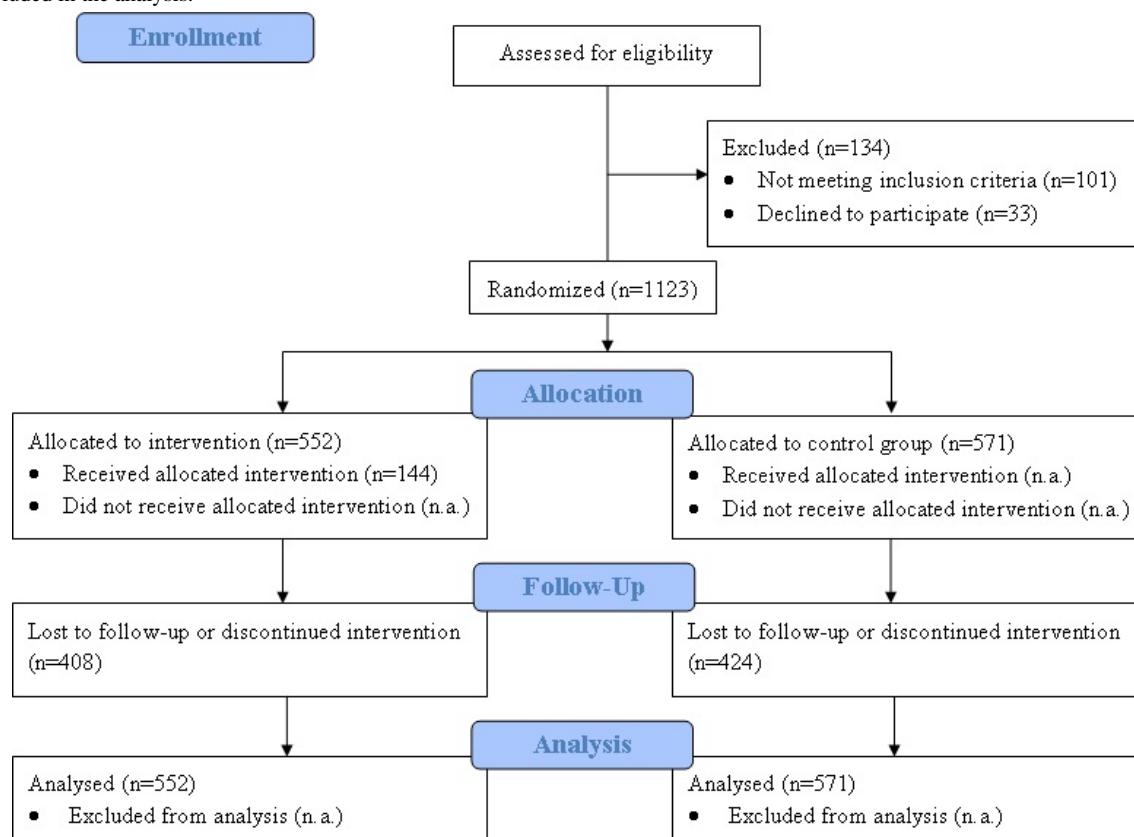
^a Fagerström Test for Nicotine Dependence.

As Table 2 shows, no differences were found with regard to baseline characteristics between respondents followed up and respondents lost to follow-up after a 6-week period. After 6

months, however, respondents lost to follow-up were significantly younger ($P = .02$) and significantly more addicted ($P = .01$) than those who remained in the study.

Table 2. Comparison between respondents followed up and respondents lost to follow-up after 6 weeks and 6 months.

Characteristic	6-week follow-up		6-month follow-up	
	Followed up (n = 449)	Lost to follow-up (n = 674)	Followed up (n = 291)	Lost to follow-up (n = 832)
Age (years), mean (SD)	50.1 (12.2)	49.4 (12.6)	50.0 (12.2)*	48.1 (12.3)*
Male, % (n)	44.5% (200)	49.7% (335)	45.5% (133)	48.4% (402)
In experimental condition, % (n)	49.9% (224)	48.7% (328)	49.3% (144)	49.1% (408)
Educational level, % (n)				
High	19% (85)	22.7% (153)	19% (54)	22.1% (184)
Medium	45.9% (206)	45.5% (307)	45.2% (132)	45.8% (381)
Low	35.2% (158)	31.8% (214)	36.3% (106)	32.0% (266)
Dutch, % (n)	98.2% (441)	97.3% (656)	97.3% (284)	97.8% (813)
With cardiovascular diseases, % (n)	12% (52)	8% (54)	11% (31)	9% (75)
With respiratory diseases, % (n)	15% (68)	14% (93)	17% (49)	13.5% (112)
Number of cigarettes smoked/day, mean (SD)	19.8 (12.1)	17.8 (6.1)	19.5 (11.4)	21.0 (12.7)
FTND ^a score (range 1–10), mean (SD)	4.8 (2.3)	4.6 (2.3)	4.7 (2.3)*	5.2 (2.5)*
Number of previous quit attempts, mean (SD)	5.0 (10.6)	5.5 (5.9)	5.1 (10.0)	5.6 (19.5)

^a Fagerström Test for Nicotine Dependence.* $P < .05$.**Figure 3.** Flow of respondents from enrollment in the study to allocation to the experimental and control conditions, retention, and whether they were included in the analysis.

Effect of the Intervention on Abstinence

Of the 552 respondents in the intervention group, 91 (17%) reported that they had refrained from smoking during the past 24 hours, 74 (13%) reported that they had not smoked during

the past 7 days, and 60 (11%) reported that they had not smoked since the previous measurement 2 days after their quit date. In the control group (n = 571) these numbers were 55 (10%), 38 (7%), and 33 (6%), respectively. The intervention had a

significant effect on all outcome measures, even when controlling for the baseline difference between the intervention and control groups with regard to their level of education (Table 3). Significantly more respondents in the intervention group than in the control group reported having been abstinent for the past 24 hours, the past 7 days, or since the previous measurement. Results from complete-case analyses were similar (Multimedia Appendix 1).

After 6 months, a total of 51 (9%) respondents in the intervention group reported having refrained from smoking during the past 24 hours, 45 (8%) reported not having smoked during the past 7 days, and 23 (4%) reported not having smoked

since the previous measurement. In the control group these numbers were 36 (6%), 34 (6%), and 19 (3%), respectively. Table 4 shows that no significant intervention effects were found with regard to all outcome measures reported at the 6-month follow-up. The complete-case analyses yielded similar results, though they were slightly more positive regarding 24-hour point prevalence abstinence (Multimedia Appendix 1).

We investigated interaction effects between condition and baseline demographic or behavioral measures, although none of these turned out to have a significant influence on any of the abstinence measures reported after 6 weeks or 6 months (data not reported).

Table 3. Effects of the Web-based smoking cessation intervention on several behavioral outcomes at 6-week follow-up among Dutch adult smokers (N = 1123) recruited from December 2009 to June 2010.

Model	24-hour ppa ^a			7-day ppa			Prolonged abstinence		
	OR ^b	95% CI ^c	P value	OR	95% CI	P value	OR	95% CI	P value
Intervention ^d	1.85	1.30–2.65	.001*	2.17	1.44–3.27	<.001*	1.99	1.28–3.09	.002*
Intervention^d	1.81	1.26–2.59	.001*	2.16	1.43–3.25	<.001*	1.96	1.26–3.05	.003*
Medium education ^e	0.81	0.51–1.32	.42	0.75	0.45–1.36	.28	0.75	0.41–1.32	.31
High education ^e	1.29	0.81–2.08	.29	0.97	0.58–1.64	.91	1.08	0.61–1.90	.80

^a Point prevalence abstinence.

^b Odds ratio.

^c Confidence interval.

^d Control group is the reference category.

^e Low education is the reference category.

*P < .05.

Table 4. Effects of the Web-based smoking cessation intervention on several behavioral outcomes at 6-month follow-up among Dutch adult smokers (N = 1123) recruited from December 2009 to June 2010.

Model	24-hour ppa ^a			7-day ppa			Prolonged abstinence		
	OR ^b	95% CI ^c	P value	OR	95% CI	P value	OR	95% CI	P value
Intervention ^d	1.51	0.97–2.35	.07	1.40	0.88–2.22	.16	1.26	0.68–2.34	.46
Intervention^d	1.47	0.94–2.30	.09	1.38	0.87–2.20	.17	1.29	0.69–2.41	.42
Medium education ^e	0.88	0.48–1.62	.69	0.86	0.47–1.58	.62	0.59	0.28–1.24	.16
High education ^e	1.38	0.76–2.52	.29	1.10	0.59–2.05	.76	0.56	0.25–1.26	.16

^a Point prevalence abstinence.

^b Odds ratio.

^c Confidence interval.

^d Control group is the reference category.

^e Low education is the reference category.

Discussion

Main Findings

In the present study we investigated the effects of a multiple computer-tailored smoking cessation program delivered through

the Internet. The results presented suggest significant effects of the intervention on short-term abstinence: at the 6-week follow-up, respondents who received the intervention were more likely to report being abstinent for the past 24 hours, for the past 7 days, and since the previous measurement (ie, 2 days after their quit date) than those who did not receive the

intervention. Despite incorporating goal and relapse prevention strategies (action and coping plans), however, we found no effect of the intervention on abstinence measures assessed after 6 months.

A potential explanation for not finding any suggestion of intervention effects on long-term abstinence might be that more than 70% of the values on the primary outcome measures had to be replaced, as our study had relatively high levels of attrition, as have many previously developed Web-based interventions [16,33-37]. In line with the Russell standard [37], we chose a negative scenario to replace missing values (ie, we considered respondents lost to follow-up to still be smoking). Although this is a recommended analysis when having to deal with dropout [37], replacing more than 70% of the values on the primary outcome measures might increase the chances of making a type II error [34,38]. This may have resulted in an underestimation of the intervention's effectiveness for those who continued to use it [34]. On the other hand, analyses with complete cases only are likely to increase the chances of a type I error, which would have resulted in unjustified conclusions in favor of the intervention studied. Although in our study the results from both types of analyses were congruent with each other, the possibilities of errors should be kept in mind when interpreting the results presented. To prevent these problems from occurring in the first place, it is of utmost importance to identify strategies that will ensure the sustained use of Web-based interventions. Previously, several suggestions have been made to prevent attrition, such as ensuring high levels of motivation to quit, providing prompts or reminders, preventing self-control depletion, for example, by having respondents form implementation intentions [36] and providing incentives of at least €10 [39]. In addition, a recent review showed that interventions that combined several of these strategies were most effective at facilitating exposure to Web-based interventions [40]. While we took all of the strategies mentioned into account, attrition rates in this study remained high. Evidently, more research is needed to identify strategies that will prevent smokers from dropping out of Web-based behavior change interventions. Qualitative research among respondents lost to follow-up might further illuminate the main reasons why these respondents discontinued a Web-based intervention. In a recent study conducted among problem drinkers, the most common reasons for not completing a Web-based intervention were personal reasons unrelated to the Web-based intervention, followed by dissatisfaction with the intervention and satisfaction with the improvement in their condition [41]. Based on the identification of the reasons for discontinuation, novel strategies to prevent attrition need to be developed and tested. In addition, more research is needed to identify strategies for dealing with missing data due to high attrition rates. A recent study comparing six different approaches to missing data concluded that multiple imputation might yield the most valid results [33]. However, as the assumption that respondents who drop out should be considered to still be smoking is well established and a still-recommended strategy in smoking cessation research [37], in the present study we opted for this strategy.

Another possible explanation for the lack of intervention effects on long-term abstinence may be that Web-based smoking

cessation programs are not sufficiently tailored and adapted to the long-term wishes of recent ex-smokers to prevent relapse to smoking. Respondents received feedback only at fixed points in time; it was not possible to obtain additional personal feedback or support at times when smokers might have needed it most. The integration of ecological momentary assessment, by collecting real-time data through, for example, palmtops, personal digital assistants, or electronic diaries, might be promising. Studies using palmtop computers showed that a decrease in self-efficacy, an increase in positive smoking outcome expectancies, and an increase in negative affect predicted the occurrence of a lapse to smoking on the next day [42,43]. Integrating ecological momentary assessment into a Web-based intervention might enable us to monitor fluctuations in factors such as self-efficacy and negative affect and, as a consequence, enable us to adapt intervention materials to the needs of recent ex-smokers and, ultimately, to prevent lapses and relapse. In addition, the finding that those lost to follow-up were significantly more addicted to nicotine than those who remained in the study supports the idea that insufficient attention was paid to dealing with withdrawal symptoms. Although, in line with current guidelines [44], we advised smokers who reported smoking more than 10 cigarettes per day to use smoking cessation medication, we did not assess whether these smokers did in fact use such medication during their quit attempt. Even though the Web-based intervention provided information on physical withdrawal symptoms and how to deal with these symptoms, all feedback messages targeted cognitions. As a consequence, solely reading these messages might not have decreased physical withdrawal symptoms. As addiction has been shown to be the most important predictor of a quit attempt's success [45,46], it may be possible to obtain higher success rates when Web-based smoking cessation interventions are combined with smoking cessation medication aimed at reducing physical withdrawal symptoms. Varenicline, for instance, has been shown to attenuate physical withdrawal symptoms and to prevent relapse to smoking [47,48].

We found no support for different intervention effects for specific subgroups of smokers. Based on the results, it could thus be argued that the intervention was equally effective for all smokers who participated in the program. However, respondents who dropped out of the study were relatively more addicted and relatively younger than those who remained in the study, which is in line with previous research [49,50]. A potential explanation might be that younger people have not yet experienced any smoking-related health effects and are, compared with older people who are more often confronted with chronic diseases, less internally motivated to invest time in health behavior change interventions [51].

Study Strengths and Limitations

Major strengths of the present study were the large sample of smokers who initiated participation in the smoking cessation program and the relatively long follow-up period. However, as mentioned previously, the study had relatively high dropout rates. In the present study, we applied several strategies previously suggested to prevent attrition [36,39]: using motivation to quit as an inclusion criterion, sending two reminder emails for each follow-up questionnaire, encouraging

respondents to formulate coping plans in the form of implementation intentions, and providing respondents with a €10 voucher for completing all follow-up questionnaires. Despite the actions taken, however, attrition rates remained high. A second limitation is that we could conduct no appropriate dose–response analysis. Insufficient data were available for participants who received one, two, or three letters and who also provided 6-month follow-up data. Of the respondents in the intervention group who provided 6-month follow-up data ($n = 144$), almost 80% ($n = 115$) received the highest dose of three feedback letters, which resulted in insufficient variation in the doses received to conduct this analysis. As previously stated, we found a significant dose–response relationship between the number of feedback moments and smoking abstinence [13]. We therefore recommend that future studies conduct a dose–response analysis to determine whether this effect can be replicated. Finally, we were unable to use continued abstinence as an outcome measure, as all respondents were asked to set a quit date within 4 weeks

from filling out the baseline questionnaire and were not obliged to quit immediately. According to the Russell standard, however, continued abstinence may classify too many successes as failures due to its strict criteria [37]. As we used prolonged abstinence instead, we were still able to assess a long period of abstinence (ie, at least 4.5 months).

Conclusions

This Web-based computer-tailored smoking cessation program had a significant effect on abstinence measured after a 6-week follow-up period. However, this effect had entirely disappeared after 6 months. To prevent relapse, future studies should focus on the possibility of applying an ecological momentary assessment or combining the present Web-based intervention with the use of smoking cessation medication. Moreover, further research should aim at identifying strategies to prevent smokers from dropping out of Web-based smoking cessation interventions. As complete-case analyses and the replacement of missing values using a negative scenario both have their limitations, alternative strategies should be identified and tested.

Acknowledgments

This study was funded by the Dutch Cancer Society (UM 2007-3834).

Conflicts of Interest

Hein de Vries is scientific director of Vision2Health, a company that licenses evidence-based innovative computer-tailored health communication tools.

Multimedia Appendix 1

Complete-case analyses.

[PDF File (Adobe PDF File), 54KB - [jmir_v14i3e82_app1.pdf](#)]

Multimedia Appendix 2

CONSORT-EHEALTH Checklist (V1.6) [52].

[PDF File (Adobe PDF File), 582KB - [jmir_v14i3e82_app2.pdf](#)]

References

1. WHO Report on the Global Tobacco Epidemic, 2008: the MPOWER package. Geneva: World Health Organization; 2008.
2. Borland R, Balmford J, Hunt D. The effectiveness of personally tailored computer-generated advice letters for smoking cessation. *Addiction* 2004 Mar;99(3):369-377. [doi: [10.1111/j.1360-0443.2003.00623.x](#)] [Medline: [14982550](#)]
3. Lancaster T, Stead LF. Self-help interventions for smoking cessation. *Cochrane Database Syst Rev* 2005(3):CD001118. [doi: [10.1002/14651858.CD001118.pub2](#)] [Medline: [16034855](#)]
4. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007 Jul;133(4):673-693. [doi: [10.1037/0033-2909.133.4.673](#)] [Medline: [17592961](#)]
5. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. *Patient Educ Couns* 1999 Feb;36(2):107-117. [Medline: [10223016](#)]
6. Dijkstra A, De Vries H, Roijackers J. Long-term effectiveness of computer-generated tailored feedback in smoking cessation. *Health Educ Res* 1998 Jun;13(2):207-214 [FREE Full text] [Medline: [10181019](#)]
7. Hoving C. The Feasibility of a Smoking Cessation Computer Tailored Expert System in Dutch General Practice and Community Pharmacy Setting [dissertation]. Maastricht: Maastricht University; 2007.
8. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 1999.
9. de Vries H, Brug J. Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach. *Patient Educ Couns* 1999 Feb;36(2):99-105. [Medline: [10223015](#)]

10. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. *Health Educ Res* 2005 Oct;20(5):527-539 [FREE Full text] [doi: [10.1093/her/cyh014](https://doi.org/10.1093/her/cyh014)] [Medline: [15701665](#)]
11. Te Poel F, Bolman C, Reubsat A, de Vries H. Efficacy of a single computer-tailored e-mail for smoking cessation: results after 6 months. *Health Educ Res* 2009 Dec;24(6):930-940 [FREE Full text] [doi: [10.1093/her/cyp036](https://doi.org/10.1093/her/cyp036)] [Medline: [19574405](#)]
12. Dijkstra A, De Vries H, Roijackers J, van Breukelen G. Tailoring information to enhance quitting in smokers with low motivation to quit: three basic efficacy questions. *Health Psychol* 1998 Nov;17(6):513-519. [Medline: [9848801](#)]
13. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med* 2010;51(3-4):214-221. [doi: [10.1016/j.ypmed.2010.06.004](https://doi.org/10.1016/j.ypmed.2010.06.004)] [Medline: [20558196](#)]
14. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. *J Med Internet Res* 2006;8(2):e10 [FREE Full text] [doi: [10.2196/jmir.8.2.e10](https://doi.org/10.2196/jmir.8.2.e10)] [Medline: [16867965](#)]
15. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. *Patient Educ Couns* 2009 Feb;74(2):156-173. [doi: [10.1016/j.pec.2008.08.023](https://doi.org/10.1016/j.pec.2008.08.023)] [Medline: [18947966](#)]
16. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. *Addiction* 2009 Nov;104(11):1792-1804. [doi: [10.1111/j.1360-0443.2009.02710.x](https://doi.org/10.1111/j.1360-0443.2009.02710.x)] [Medline: [19832783](#)]
17. Centraal Bureau voor de Statistiek. 2008 Dec 17. 1,2 Miljoen Nederlanders Hebben Thuis Geen Internet URL: <http://www.cbs.nl/nl-NL/menu/themas/vrije-tijd-cultuur/publicaties/artikelen/archief/2008/2008-2641-wm.htm> [accessed 2011-03-21] [WebCite Cache ID 5xLwWmphS]
18. Internet World Stats. Miniwatts Marketing Group. 2010. Internet Usage Statistics: The Internet Big Picture URL: <http://www.internetworldstats.com/stats.htm> [accessed 2011-03-21] [WebCite Cache ID 5xLwg1Msm]
19. Dijkstra A, De Vries H. The development of computer-generated tailored interventions. *Patient Educ Couns* 1999 Feb;36(2):193-203. [Medline: [10223023](#)]
20. Hughes JR, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction* 2004 Jan;99(1):29-38. [Medline: [14678060](#)]
21. Rabois D, Haaga DA. The influence of cognitive coping and mood on smokers' self-efficacy and temptation. *Addict Behav* 2003 Apr;28(3):561-573. [Medline: [12628627](#)]
22. Leventhal AM, Ramsey SE, Brown RA, LaChance HR, Kahler CW. Dimensions of depressive symptoms and smoking cessation. *Nicotine Tob Res* 2008 Mar;10(3):507-517. [doi: [10.1080/14622200801901971](https://doi.org/10.1080/14622200801901971)] [Medline: [18324570](#)]
23. Fucito LM, Latimer AE, Carlin-Menter S, Salovey P, Cummings KM, Makuch RW, et al. Nicotine dependence as a moderator of a quitline-based message framing intervention. *Drug Alcohol Depend* 2011 Apr 1;114(2-3):229-232. [doi: [10.1016/j.drugalcdep.2010.09.012](https://doi.org/10.1016/j.drugalcdep.2010.09.012)] [Medline: [21036492](#)]
24. Strecher VJ, Shiffman S, West R. Moderators and mediators of a web-based computer-tailored smoking cessation program among nicotine patch users. *Nicotine Tob Res* 2006 Dec;8 Suppl 1:S95-101. [Medline: [17491176](#)]
25. de Vries H, Mudde A, Leijts I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. *Health Educ Res* 2003 Oct;18(5):611-626 [FREE Full text] [Medline: [14572020](#)]
26. van Osch L, Lechner L, Reubsat A, Wigger S, de Vries H. Relapse prevention in a national smoking cessation contest: effects of coping planning. *Br J Health Psychol* 2008 Sep;13(Pt 3):525-535. [doi: [10.1348/135910707X224504](https://doi.org/10.1348/135910707X224504)] [Medline: [17631695](#)]
27. Smit ES, de Vries H, Hoving C. The PAS study: a randomized controlled trial evaluating the effectiveness of a web-based multiple tailored smoking cessation programme and tailored counselling by practice nurses. *Contemp Clin Trials* 2010 May;31(3):251-258. [doi: [10.1016/j.cct.2010.03.001](https://doi.org/10.1016/j.cct.2010.03.001)] [Medline: [20215047](#)]
28. Faul F, Erdfelder E. GPOWER: A Priori, Post Hoc, and Compromise Power Analyses for MS-DOS. Bonn: Department of Psychology, University of Bonn; 1992.
29. Hoving C, Mudde AN, Dijk F, de Vries H. Effectiveness of a smoking cessation intervention in Dutch pharmacies and general practices. *Health Educ* 2010;110(1):17-29. [doi: [10.1108/09654281011008726](https://doi.org/10.1108/09654281011008726)]
30. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol* 1983 Jun;51(3):390-395. [Medline: [6863699](#)]
31. Mudde AN, Willemsen MC, Kremers S, de Vries H. Meetinstrumenten: Voor Onderzoek Naar Roken en Stoppen met Roken. Den Haag: STIVORO voor een rookvrije toekomst; 2006.
32. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br J Addict* 1991 Sep;86(9):1119-1127. [Medline: [1932883](#)]
33. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. *J Med Internet Res* 2010;12(5):e54 [FREE Full text] [doi: [10.2196/jmir.1448](https://doi.org/10.2196/jmir.1448)] [Medline: [21169167](#)]
34. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](#)]

35. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. *J Med Internet Res* 2008;10(5):e40 [FREE Full text] [doi: [10.2196/jmir.993](https://doi.org/10.2196/jmir.993)] [Medline: [19017582](https://pubmed.ncbi.nlm.nih.gov/19017582/)]
36. Webb TL. Commentary on Shahab & McEwen (2009): Understanding and preventing attrition in online smoking cessation interventions: a self-regulatory perspective. *Addiction* 2009 Nov;104(11):1805-1806. [doi: [10.1111/j.1360-0443.2009.02751.x](https://doi.org/10.1111/j.1360-0443.2009.02751.x)] [Medline: [19832784](https://pubmed.ncbi.nlm.nih.gov/19832784/)]
37. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction* 2005 Mar;100(3):299-303. [doi: [10.1111/j.1360-0443.2004.00995.x](https://doi.org/10.1111/j.1360-0443.2004.00995.x)] [Medline: [15733243](https://pubmed.ncbi.nlm.nih.gov/15733243/)]
38. Field A. *Discovering Statistics Using SPSS*. 2nd edition. London: Sage Publications; 2005.
39. Khadjesari Z, Murray E, Kalaitzaki E, White IR, McCambridge J, Thompson SG, et al. Impact and costs of incentives to reduce attrition in online trials: two randomized controlled trials. *J Med Internet Res* 2011;13(1):e26 [FREE Full text] [doi: [10.2196/jmir.1523](https://doi.org/10.2196/jmir.1523)] [Medline: [21371988](https://pubmed.ncbi.nlm.nih.gov/21371988/)]
40. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. *Health Educ Behav* 2011 Feb;38(1):49-62. [doi: [10.1177/1090198110372878](https://doi.org/10.1177/1090198110372878)] [Medline: [21189422](https://pubmed.ncbi.nlm.nih.gov/21189422/)]
41. Postel MG, de Haan HA, ter Huurne ED, van der Palen J, Becker ES, de Jong CA. Attrition in web-based treatment for problem drinkers. *J Med Internet Res* 2011;13(4):e117 [FREE Full text] [doi: [10.2196/jmir.1811](https://doi.org/10.2196/jmir.1811)] [Medline: [22201703](https://pubmed.ncbi.nlm.nih.gov/22201703/)]
42. Gwaltney CJ, Shiffman S, Balabanis MH, Paty JA. Dynamic self-efficacy and outcome expectancies: prediction of smoking lapse and relapse. *J Abnorm Psychol* 2005 Nov;114(4):661-675. [doi: [10.1037/0021-843X.114.4.661](https://doi.org/10.1037/0021-843X.114.4.661)] [Medline: [16351387](https://pubmed.ncbi.nlm.nih.gov/16351387/)]
43. Shiffman S, Balabanis MH, Gwaltney CJ, Paty JA, Gnys M, Kassel JD, et al. Prediction of lapse from associations between smoking and situational antecedents assessed by ecological momentary assessment. *Drug Alcohol Depend* 2007 Dec 1;91(2-3):159-168. [doi: [10.1016/j.drugalcdep.2007.05.017](https://doi.org/10.1016/j.drugalcdep.2007.05.017)] [Medline: [17628353](https://pubmed.ncbi.nlm.nih.gov/17628353/)]
44. Chavannes NH, Kaper J, Frijling BD, Van der Laan JR, Jansen PWM, Guerrouj S, Drenthen AJM, Bax W, Wind LA. NHG-Standaard Stoppen met roken. *Huisarts Wetenschap* 2007;50(7):306-314. [doi: [10.1007/978-90-313-6614-9_90](https://doi.org/10.1007/978-90-313-6614-9_90)]
45. McEwen A, West R, McRobbie H. Motives for smoking and their correlates in clients attending Stop Smoking treatment services. *Nicotine Tob Res* 2008 May;10(5):843-850. [doi: [10.1080/14622200802027248](https://doi.org/10.1080/14622200802027248)] [Medline: [18569758](https://pubmed.ncbi.nlm.nih.gov/18569758/)]
46. West R. The multiple facets of cigarette addiction and what they mean for encouraging and helping smokers to stop. *COPD* 2009 Aug;6(4):277-283. [Medline: [19811387](https://pubmed.ncbi.nlm.nih.gov/19811387/)]
47. Hajek P, Stead LF, West R, Jarvis M, Lancaster T. Relapse prevention interventions for smoking cessation. *Cochrane Database Syst Rev* 2009(1):CD003999. [doi: [10.1002/14651858.CD003999.pub3](https://doi.org/10.1002/14651858.CD003999.pub3)] [Medline: [19160228](https://pubmed.ncbi.nlm.nih.gov/19160228/)]
48. Keating GM, Siddiqui MA. Varenicline: a review of its use as an aid to smoking cessation therapy. *CNS Drugs* 2006;20(11):945-960. [Medline: [17044731](https://pubmed.ncbi.nlm.nih.gov/17044731/)]
49. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. *J Med Internet Res* 2008;10(5):e36 [FREE Full text] [doi: [10.2196/jmir.1002](https://doi.org/10.2196/jmir.1002)] [Medline: [18984557](https://pubmed.ncbi.nlm.nih.gov/18984557/)]
50. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. *Patient Prefer Adherence* 2008;2:57-65 [FREE Full text] [Medline: [19920945](https://pubmed.ncbi.nlm.nih.gov/19920945/)]
51. de Nooijer J, Oenema A, Klok G, Brug H, De Vries H, De Vries N. *Bevordering van Gezond Gedrag via Internet: Nu en in de Toekomst*. Maastricht: Maastricht University; 2005.
52. 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](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]

Abbreviations

CI: confidence interval

OR: odds ratio

Edited by G Eysenbach; submitted 21.03.11; peer-reviewed by L Shahab, J McClure, J Balmford; comments to author 13.05.11; revised version received 05.07.11; accepted 12.03.12; published 11.06.12.

Please cite as:

Smit ES, de Vries H, Hoving C

Effectiveness of a Web-Based Multiple Tailored Smoking Cessation Program: A Randomized Controlled Trial Among Dutch Adult Smokers

J Med Internet Res 2012;14(3):e82

URL: <http://www.jmir.org/2012/3/e82/>

doi: [10.2196/jmir.1812](https://doi.org/10.2196/jmir.1812)

PMID: [22687887](https://pubmed.ncbi.nlm.nih.gov/22687887/)

©Eline Suzanne Smit, Hein de Vries, Ciska Hoving. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 11.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Efficacy of a Computerized Simulation in Promoting Walking in Individuals With Diabetes

Bryan Gibson^{1,2}, DPT, PhD; Robin L Marcus³, PT, PhD; Nancy Staggers⁴, RN, PhD; Jason Jones⁵, PhD; Matthew Samore^{1,2}, MD; Charlene Weir^{1,2}, RN, PhD

¹George E Whalen Salt Lake City Veterans Affairs Healthcare System, IDEAS center, Salt Lake City, UT, United States

²Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

³Department of Physical Therapy, University of Utah, Salt Lake City, UT, United States

⁴Department of Nursing Informatics, University of Maryland, Baltimore, MD, United States

⁵Kaiser Permanente Southern California, Pasadena, CA, United States

Corresponding Author:

Bryan Gibson, DPT, PhD

George E Whalen Salt Lake City Veterans Affairs Healthcare System

IDEAS center

500 Foothill drive

Salt Lake City, UT, 84148

United States

Phone: 1 (801) 582 1565 ext 246

Fax: 1 (801) 584 5640

Email: bryan.gibson@utah.edu

Abstract

Background: Regular walking is a recommended but underused self-management strategy for individuals with type 2 diabetes mellitus (T2DM).

Objective: To test the impact of a simulation-based intervention on the beliefs, intentions, knowledge, and walking behavior of individuals with T2DM. We compared two versions of a brief narrated simulation. The experimental manipulation included two components: the presentation of the expected effect of walking on the glucose curve; and the completion of an action plan for walking over the next week. Primary hypotheses were (1) intervention participants' walking (minutes/week) would increase more than control participants' walking, and (2) change in outcome expectancies (beliefs) would be a function of the discrepancy between prior beliefs and those presented in the simulation. Secondary hypotheses were that, overall, behavioral intentions to walk in the coming week and diabetes-related knowledge would increase in both groups.

Methods: Individuals were randomly assigned to condition. Preintervention measures included self-reported physical activity (International Physical Activity Questionnaire [IPAQ] 7-day), theory of planned behavior-related beliefs, and knowledge (Diabetes Knowledge Test). During the narrated simulation we measured individuals' outcome expectancies regarding the effect of exercise on glucose with a novel drawing task. Postsimulation measures included theory of planned behavior beliefs, knowledge, and qualitative impressions of the narrated simulation. The IPAQ 7-day was readministered by phone 1 week later. We used a linear model that accounted for baseline walking to test the main hypothesis regarding walking. Discrepancy scores were calculated between the presented outcome and individuals' prior expectations (measured by the drawing task). A linear model with an interaction between intervention status and the discrepancy score was used to test the hypothesis regarding change in outcome expectancy. Pre-post changes in intention and knowledge were tested using paired *t* tests.

Results: Of 65 participants, 33 were in the intervention group and 32 in the control group. We excluded 2 participants from analysis due to being extreme outliers in baseline walking. After adjustment for baseline difference in age and intentions between groups, intervention participants increased walking by 61.0 minutes/week (SE 30.5, $t_{58} = 1.9$, $P = .05$) more than controls. The proposed interaction between the presented outcome and the individual's prior beliefs was supported: after adjustment for baseline differences in age and intentions between groups, the coefficient for the interaction was $-.25$, (SE 0.07, $t_{57} = -3.2$, $P < .01$). On average participants in both groups improved significantly from baseline in intentions (mean difference 0.66, $t_{62} = 4.5$, $P < .001$) and knowledge (mean difference 0.38, $t_{62} = 2.4$, $P = .02$).

Conclusions: This study suggests that a brief, Internet-ready, simulation-based intervention can improve knowledge, beliefs, intentions, and short-term behavior in individuals with T2DM.

(*J Med Internet Res* 2012;14(3):e71) doi:[10.2196/jmir.1965](https://doi.org/10.2196/jmir.1965)

KEYWORDS

Computer simulation; type 2 diabetes mellitus; physical activity; blood glucose

Introduction

Type 2 diabetes mellitus (T2DM) affects approximately 24 million people in the United States, and is associated with significant morbidity and early mortality [1]. Regular physical activity has been shown to improve glycemic control [2,3], reduce blood pressure [4], reduce lipids [4], and improve cardiorespiratory fitness in individuals with T2DM [5]. These intermediate outcomes have been associated with diabetes-related morbidity and mortality [6]. Although physical activity is considered one of the three pillars of diabetes self-management [7], most people with T2DM do not perform sufficient amounts [8].

There are many reasons why individuals with T2DM may not perform an appropriate self-management behavior such as being active. In this study we used a brief, narrated simulation to address two factors that we believe are amenable to an informatics intervention: inaccurate mental models of the effects of behavior on the disease [9-11] and difficulties in translating good intentions into action [12].

Glucose Curves

The intervention in this study was based on simulated glucose curves. Glucose curves represent an individual's variation in plasma glucose through a day. Prior work suggests that glucose curves may be useful as an interface for educational and motivational interventions. Small trials of participants with type 1 diabetes have shown that classroom education using simulated glucose curves positively affects knowledge [13], the frequency of hypoglycemic events [14], and hemoglobin A1c [14]. In T2DM, interviews with individuals before and after viewing their own glucose curves suggest that viewing the curves appears to provide individuals with a greater understanding of the daily variation in glucose (particularly postprandial peaks) and may result in greater intention to perform self-care activities, including to be more physically active [15]. We believe glucose curves offer value because they provide contextual information that individual self-monitored glucose values do not provide.

Theory of Planned Behavior

According to the theory of planned behavior, an individual's intention to perform a behavior is a function of their beliefs. In this study we focused on a particular type of belief: *outcome expectancies*. Outcome expectancies are an individual's belief regarding the likely outcome of a given behavior. The intervention version of our simulation demonstrates the expected change in the glucose curve with both a single walk and regular walking over time.

Prior work has shown that outcome expectancies are related to self-care behaviors in individuals with T2DM [16-19] and that

individuals with T2DM generally have low outcome expectancies regarding the effect of exercise on blood glucose [19]. We are not aware of studies that have attempted to *change* outcome expectancies in this population. In general, interventions targeted at outcome expectancies related to physical activity have shown limited efficacy in most populations [20].

Implementation Intentions

While the beliefs included in the theory of planned behavior have been shown to predict the intentions of individuals with T2DM to be physically active [16], changes in behavioral intention are only moderately predictive of actual changes in behavior [21]. *Implementation intentions* are if-then plans linking specific cues in the environment to a desired behavior. Implementation intentions have been found to be strongly effective in translating intentions into action [12,22]. Recent evidence suggests that individuals who mentally simulate the behavior as they create the implementation intention are even more successful in acting on their intentions [23,24].

The intervention version of our simulation guided participants through writing an action plan for walking while concurrently mentally simulating the planned behavior. In this plan participants indicated where, when, with whom, and for how long they would walk for each day in the next week.

Our hypotheses in this trial were that (1) individuals viewing the intervention version of the narrated simulation would report more walking in the subsequent week than control participants would, and (2) changes in outcome expectancies for intervention participants would vary as a function of the discrepancy between the effect presented in the simulation and the individual's prior beliefs. Finally, we hypothesized that, overall, both groups would increase their behavioral intentions to walk in the subsequent week and their diabetes-related knowledge.

Methods

Participants

We recruited participants between March 2010 and August 2011 at the George E. Whalen Department of Veterans Affairs Medical Center (Salt Lake City, UT, USA) in primary care clinics, diabetes education and weight management classes, a biweekly diabetes exercise group at the University of Utah, a community diabetes health fair, and via an email to a diabetes-related listserv.

Our inclusion criteria were that participants be between 30 and 70 years of age, have a diagnosis of T2DM, and be able to speak English fluently. Participants with a diagnosis of dementia or severe mental disease, using insulin, or having microvascular or macrovascular complications of diabetes were excluded. The

rationale for these last two criteria was 2-fold: first, the content of the narrated simulation is geared toward individuals taking oral medications, and second, we wanted to minimize the risk of walking-induced hypoglycemia, foot ulceration, or a cardiac event. Initial recruitment efforts were exclusively among veterans at the Salt Lake City Veterans Administration Healthcare System, aged 40–60 years; however, due to slow recruitment, in June 2010 we expanded recruitment to the larger community and a wider age range.

Settings

The study was conducted in a location convenient to the participant. These locations included the Salt Lake City VA library, a room adjacent to the exercise room at the diabetes exercise group, a table at a diabetes health fair, a meeting room at a public library, and a private office. All meetings were

between the principal investigator (BG) and individual participants.

Description of the Simulation

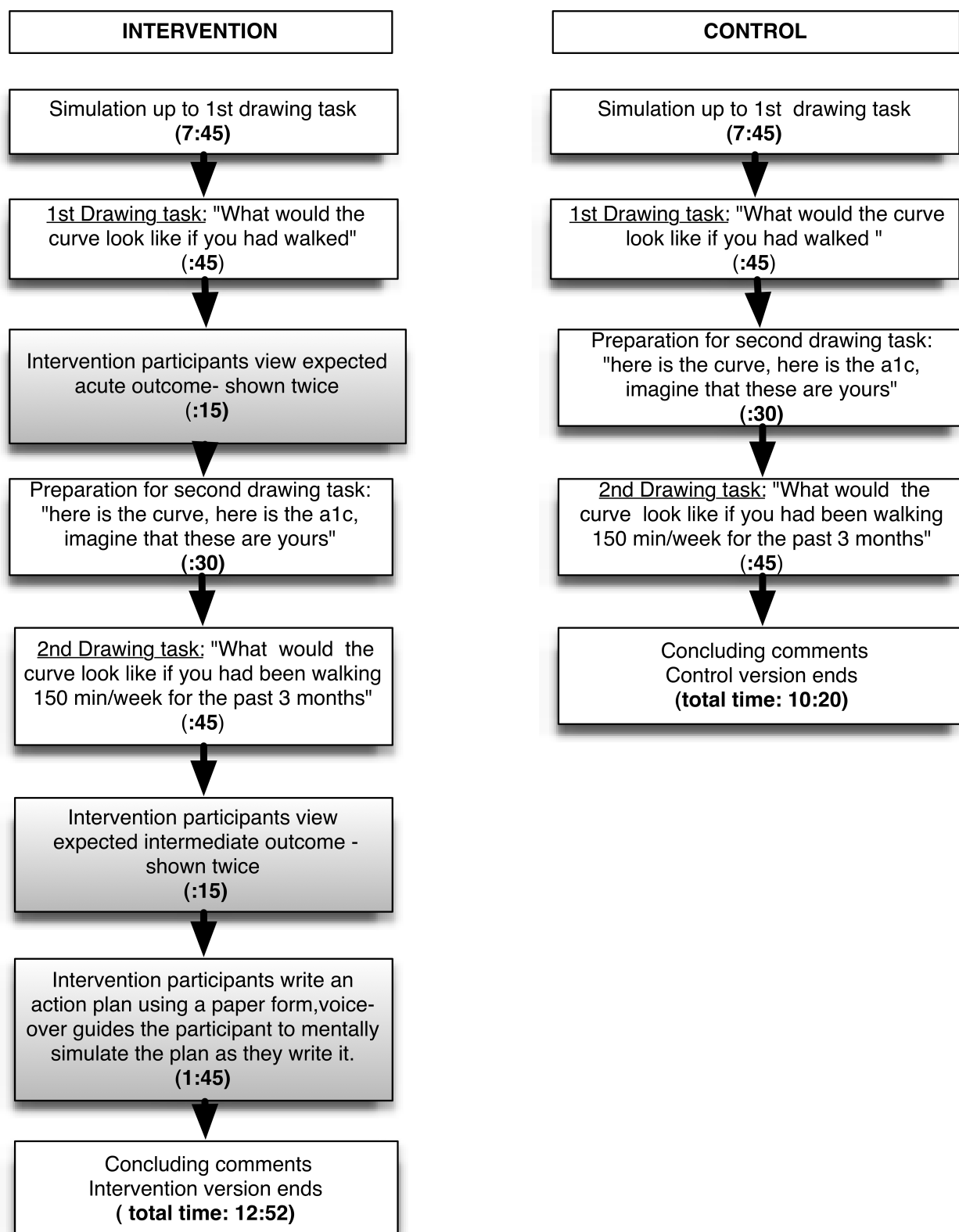
The narrated simulation is based on simulated glucose curves [25]. Concepts are presented using the curves without numbers, supplemented by simple icons. A voiceover and music soundtrack accompany the narrated simulation (see [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#) for the intervention and control simulations). [Table 1](#) lists the concepts addressed in the narrated simulation and the time used to explain each concept.

Participants were shown one of two versions of the simulation. The intervention version and the control version were identical through the first 8 minutes and 30 seconds ([Figure 1](#)).

Table 1. Concepts included in the narrated simulation and their timing.

Concept	Timing (minutes and seconds)
What is the glucose curve?	1:40
When is blood sugar highest and when is it lowest?	0:20
How do meals affect the glucose curve?	0:30
What is the dawn phenomenon?	0:30
What is the safe range of blood sugar?	0:40
What is hemoglobin A1c?	0:15
How does the blood sugar curve change (over years) as A1c increases?	1:40
Why is high blood sugar bad for you? (Includes photographs of individuals with microvascular complications)	1:40
How are changes in A1c associated with complications?	0:20
What can you do today to control your blood sugar?	0:35

Figure 1. Procedures in the simulation for the intervention and control groups. Boxes with a gray background show intervention-specific components. Duration is in minutes and seconds.



Drawing Tasks

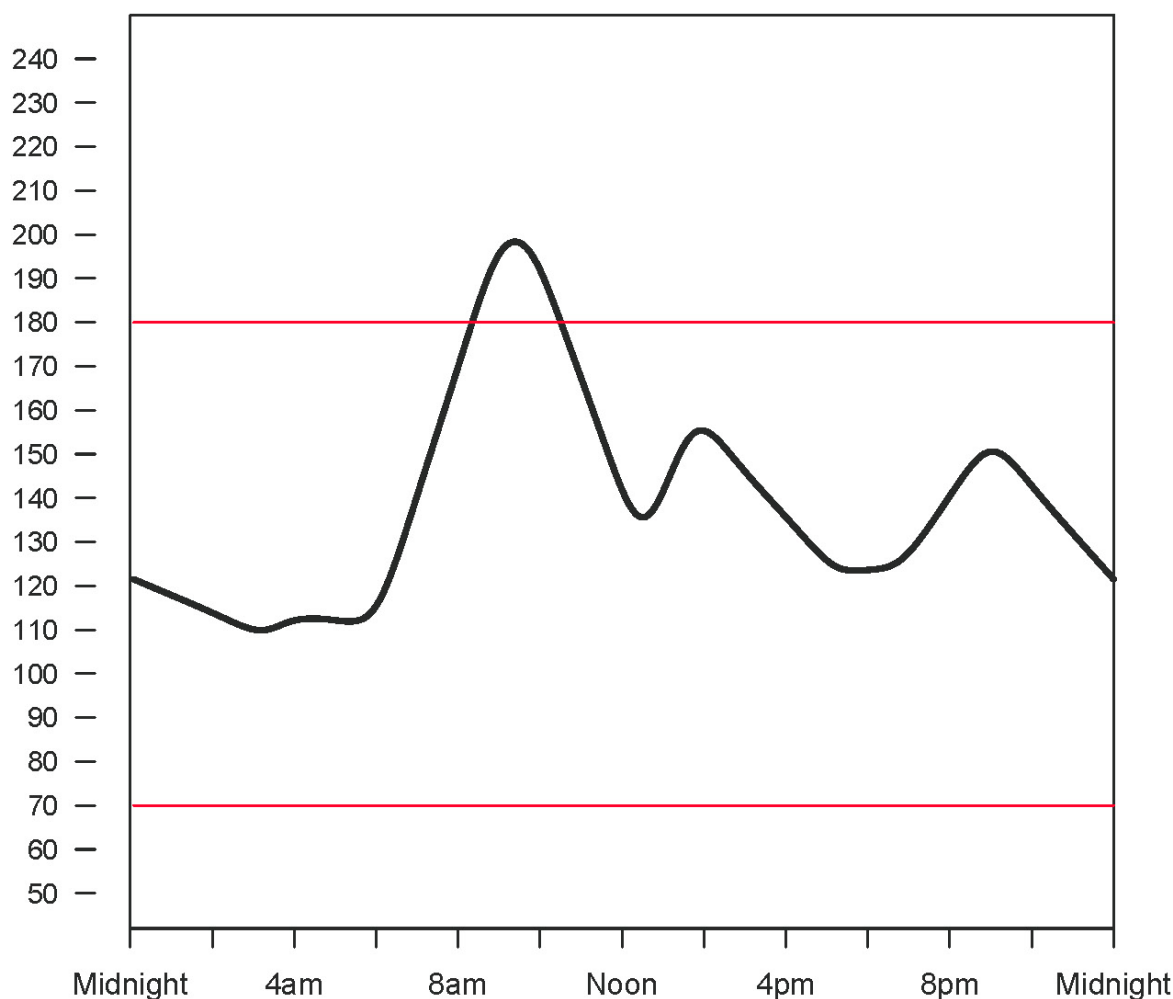
At this point in the narrated simulation, participants were shown a glucose curve of an individual "who has had diabetes for a few years," and the voiceover asked them to imagine that the curve was their glucose curve from yesterday. Using a paper

copy of the curve on the screen (Figure 2), participants were asked to draw what they thought the curve would have looked like if they had gone for a "30-minute walk yesterday an hour after breakfast." As a second drawing task participants were shown the same curve of an individual "who has had diabetes

for a few years,” asked to imagine that it was their curve from yesterday, and asked to draw what they thought the curve would have looked like if they had been walking 5 days a week for 30 minutes each for the past 3 months. The purpose of these two drawing tasks was to capture the individual’s outcome

expectancy regarding the change in glucose with a single walk and the change in hemoglobin A1c with regular walking. The advantage of this method is that it allowed us to measure the individuals’ outcome expectancy across three dimensions: the magnitude, direction, and duration of change in the curve.

Figure 2. Simulated glucose curve used in the drawing task.



Difference Between Control and Intervention Conditions

The control version of the narrated simulation ended after the two drawing tasks. In the intervention version of the narrated simulation, after completing each drawing task, viewers were shown the expected change in the curve. They were then guided by the voiceover to complete a paper plan of their walking over the next week: how many days they would walk, on which days

they would walk, how long each walk would be, in what location they would walk, at what time of day, with whom, and any preparatory actions they would take to facilitate the plan (eg, put walking shoes in their car) (Figure 3). As participants completed the paper plan, the voiceover guided them to mentally simulate the plan. These procedures were specifically designed to facilitate the formation of implementation intentions in the minds of the participants.

Figure 3. Walking plan to be completed by intervention participants.

Make a Plan for Your Walks over the Next Week

How many days in the next week do you plan to walk ? _____

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Where will you walk? (ex: In the park)							
At what time of day will you walk? (ex: 5:00 pm)							
For how long will you walk? (ex: 30 minutes)							
With whom will you walk? (ex: With my friend)							
How will you prepare for the walk? (ex: Put my walking shoes in the car)							

Motivational Components of Both Versions of the Simulation

We hypothesized that two components of the simulation might increase behavioral intentions for both groups. First, in the elicitation of individual's outcome expectancies via the drawing task, the potential outcome of exercise is framed as an upward counterfactual (how could things have been better: "what *would have* happened if you had exercised"). This was done based on theory [26] and evidence indicating that upward counterfactual thinking facilitates behavioral intentions [27]. Second, the narrated simulation presented the long-term outcomes of being sedentary to both groups ("here is how the glucose curve changes over years if you don't eat right and exercise regularly"). We included this component based on Williams and colleagues' suggestion that the construct of outcome expectancy in physical activity research should incorporate both the positive effects of increased activity and the negative effects of being sedentary [20]. We included these components because we wanted to maximize intentions in the intervention group prior to their writing the action plan for walking in the next week; this was based on prior evidence that implementation intentions are most effective when intentions are strong [28]. We did not manipulate these constructs across conditions in this study because our goal was to experimentally determine the effect of the combination of presenting potential outcomes and action planning on behavior.

Procedures

After obtaining informed consent from the participants, we collected the following measures: (1) demographic information (Table 1), (2) a 10-item version of the Diabetes Numeracy Test [29], (3) a 14-item questionnaire that measures constructs from the theory of planned behavior; this was developed in pilot testing using procedures described by Azjen [30] (see Multimedia Appendix 3 for the complete questionnaire), (4) the short telephone version of the International Physical Activity Questionnaire (IPAQ), a validated self-reported measure of physical activity over the last 7 days [31], and (5) the 14-item Diabetes Knowledge Test [32] and the 5-item ABC test [33], both tests of diabetes-specific knowledge.

Participants then watched the narrated simulation on a laptop computer while wearing headphones. During the narrated simulation, all participants completed the two drawing tasks described above. To minimize demand effects, the investigator left the room while participants watched the animation; most questionnaires were administered by paper. However, since the IPAQ was going to be readministered by phone a week later, this questionnaire was administered orally by the investigator during the in-person meeting.

After participants watched the narrated simulation, the 14-item theory of planned behavior-related questionnaire and both diabetes-related knowledge tests were repeated. In addition, to measure the degree to which participants felt that the information in the animation was personally relevant, participants answered

two 7-point Likert-type questions: “I think the glucose curves in the movie were related to *my diabetes*” and “I think the complications shown in the movie *could happen to me*.”

To conclude the in-person meeting, we asked participants about their qualitative impressions of the narrated simulation: what they liked and did not like, if there were parts of the simulation they found confusing, and if there were concepts they would like to see presented in this manner that were not included in the narrated simulation. These questions were administered orally.

We contacted participants by phone 1 week later and readministered the IPAQ measure of physical activity over the last 7 days [31]. This was followed by a questionnaire asking whether the participant thought about the glucose curves in the week since watching the narrated simulation and, if so, whether they thought about them before, during, or after eating, exercising, or testing their glucose. The purpose behind these last questions was exploratory for future work with this intervention.

Analysis

We performed all analyses using R version 2.10.0, freely available statistical computing software [34]. We excluded 2 individuals from this analysis: 1 control participant who reported walking 35 hours/week at baseline and 1 intervention participant who reported walking 18 hours/week; these individuals' baseline walking times were ≥ 2.5 standard deviations above the mean. In addition, including these individuals would have overestimated the effect of the intervention in our main hypotheses.

To test our primary hypothesis (that the intervention version of the narrated simulation would more positively affect individuals' walking), we used a linear model with intervention status and preanimation walking (minutes/week) as the covariates. We adjusted for significant between-group differences in age and a near-significant difference in baseline behavioral intent (see Table 2).

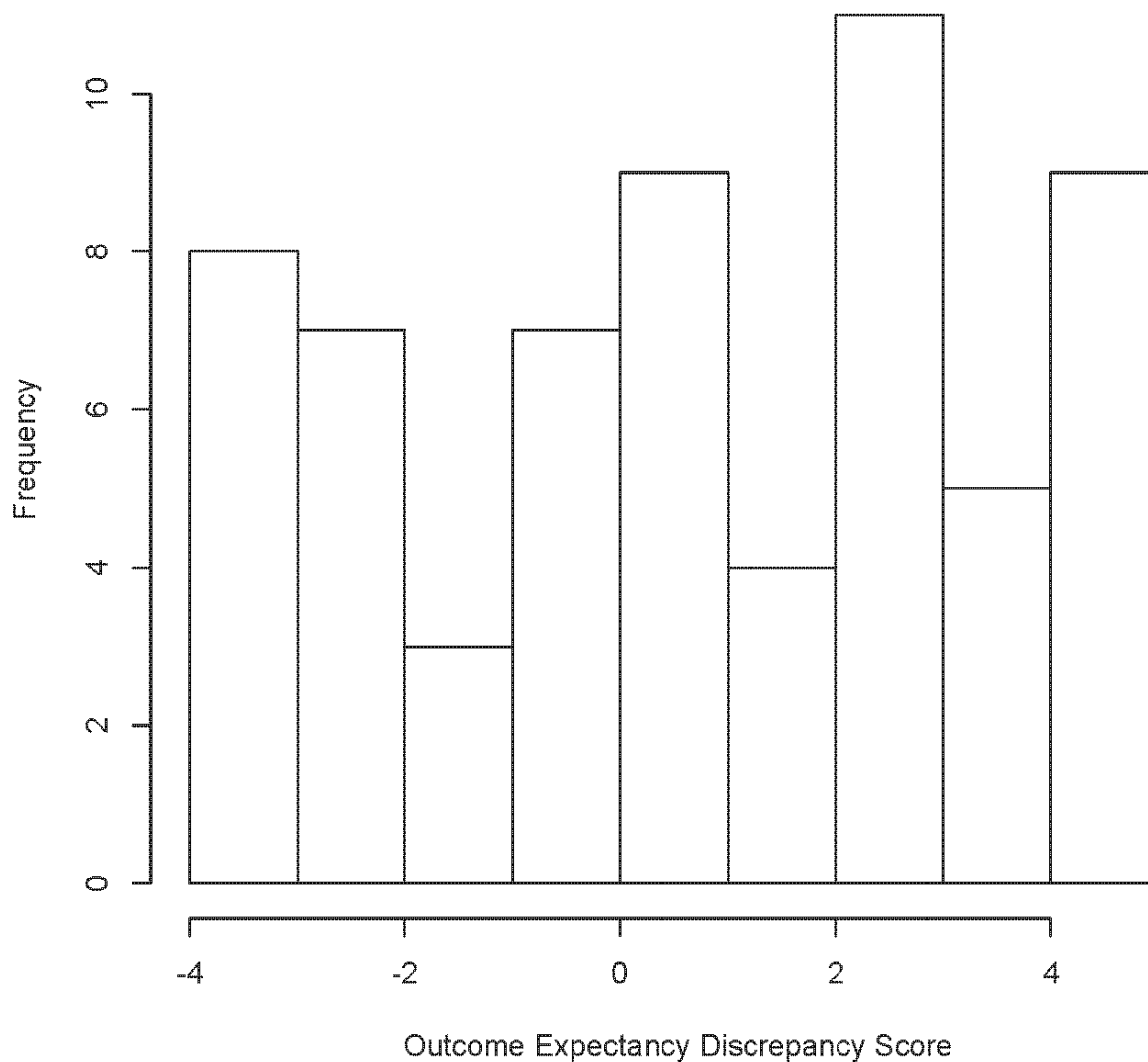
To test our second hypothesis (that among intervention participants change in outcome expectancies [beliefs] would be a function of the discrepancy between prior beliefs and those presented in the narrated simulation), we first needed to calculate the change in outcome expectancy and then calculate a score reflecting the discrepancy between the presented outcome and the individual's expected outcome. Once these scores were calculated, we used a linear model with an interaction between the discrepancy score and intervention status as a covariate after adjusting for age and baseline intent.

Outcome expectancies were measured using the following questions on the theory of planned behavior questionnaire: “Walking for at least 30 minutes will lower my blood sugar,” and “Walking for at least 30 minutes/day, 5 days a week *over the next 3 months* will lower my hemoglobin A1c.” Participants agreed or disagreed on a 7-point Likert scale (see Multimedia Appendix 3). As suggested by Azjen, for each of the pre- and post-theory of planned behavior measures, the individual's score for these two questions was averaged to reflect the overall construct of outcome expectancy [30]. A change score was calculated by subtracting the preintervention measure of outcome expectancy from the postintervention measure.

We calculated the outcome expectancy discrepancy score by measuring the difference between the presented change in the glucose curve and the individual's outcome expectancy elicited in the drawing task. We scored each dimension of the individual's outcome expectancy (direction, duration, and magnitude) according to whether the individual's outcome expectation was negative, neutral, or positive. For example, if the decrease in the individual's drawn curve was greater in magnitude than the decrease in the presented curve (positive expectancy), this dimension was scored 1. If the magnitude of the participant's expectation was the same as the presented curve, the score was 0 (accurate understanding). If the drawn magnitude was less than the presented curve, the participant was scored -1 (negative expectancy). Since the direction of the change in the curve could only increase or decrease, individuals were scored 1 if their drawing reflected a decrease (a positive expectancy and accurate understanding) and -1 if their drawing reflected an increase in blood glucose postexercise (negative expectancy). The discrepancy score used in the regression is the sum of all the dimension scores for both drawing tasks with a possible range of -6 to 6. Figure 4 is a histogram of the distribution of discrepancy scores.

To test our secondary hypotheses (that, overall, both versions of the narrated simulation would positively affect behavioral intentions and knowledge), we used paired *t* tests to compare presimulation versus postsimulation measures.

Finally, we conducted an exploratory analysis to inform future work by examining participants' responses to the qualitative questions of what they liked and did not like in the narrated simulation, what they found confusing, and what they would like to see in future versions for recurrent themes. We also examined the proportion of individuals who reported thinking about the glucose curves in the next week and the context in which they reported thinking about them.

Figure 4. Distribution of outcome expectancy discrepancy scores.

Results

Description of the Sample

Table 2 presents the baseline characteristics of the intervention and control groups. The randomization resulted in equal groups

on all measures with the exception of age; the average age of the control group was slightly higher than that of the intervention group; in addition, a near-significant difference existed in baseline intentions regarding walking in the intervention group.

Table 2. Baseline characteristics of control and intervention groups.

Characteristic	Intervention group (n = 33)	Control group (n = 32)	P value
Sex, n^a			.87
Male	20	21	
Female	13	11	
Veterans, n ^a	10	12	.72
Age (years), median (range) ^b	56 (34–70)	61 (36–70)	.02
Years since diagnosis, median (range) ^b	7 (.02–20)	8.5 (.12–19)	.96
Hemoglobin A1c, median (range) ^b	7.0 (5.6–11.8)	6.9 (6.1–10.3)	.63
Diabetes numeracy (scale of 0–10), median (range) ^b	8 (1–10)	8 (2–10)	.34
Frequency of self-monitoring (times/week), median (range) ^b	5 (0.1–21)	2.75 (0–21)	.13
Have email?, n ^a	29	29	.96
Frequency of non-job email use (x/week), median (range) ^b	14 (0–14)	14 (0–14)	.65
Have a personal health record?, n ^a	12	10	.86
Nonwalking physical activity (metabolic equivalents × minutes/week), median (range) ^b	960 (0–8820)	512 (0–8640)	.12
Walking (minutes/week), median (range) ^b	90 (0–1080)	145 (0–2100)	.27
Knowledge (Diabetes Knowledge Test, scale of 0–14), median (range) ^b	12 (5–14)	12 (6–14)	.55
Behavioral intention (scale of 1–7), median (range) ^b	5 (1–7)	6 (1–7)	.08

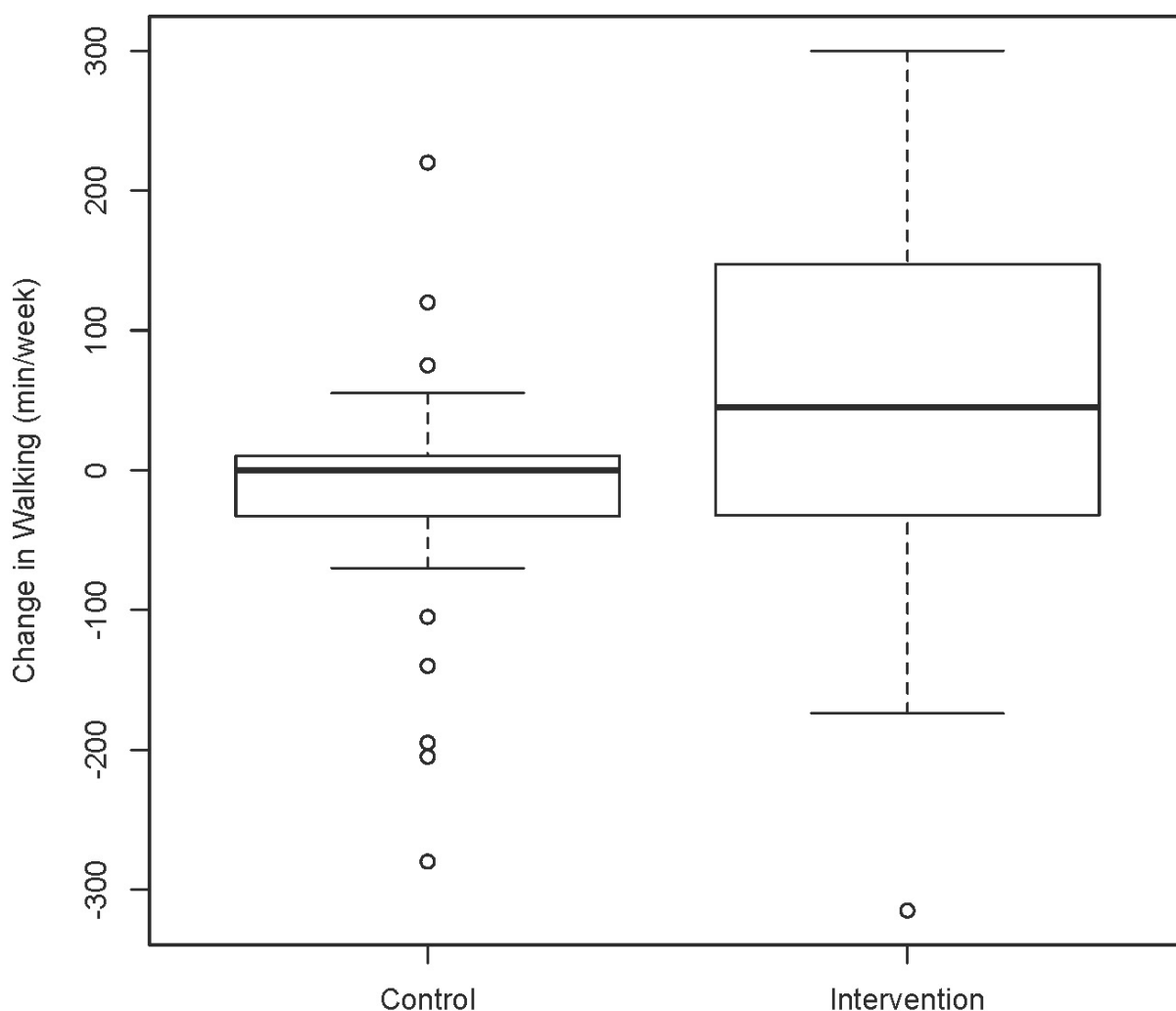
^a Chi-square test.^b Kruskal-Wallis test.

Hypothesis 1

Our first and most clinically significant hypothesis was supported: intervention participants increased walking time more than control participants. After taking into account baseline walking and adjusting for age and baseline behavioral intent,

the mean effect of the intervention was an increase of 61.0 minutes (SE 30.5, $t_{58} = 1.9$, $P = .05$). Neither age (coef = -1.2 , SE 1.9, $t_{58} = -0.6$, $P = .5$) nor baseline behavior intent (coef = 3.6, SE 9.5, $t_{58} = 0.3$, $P = .7$) was a significant predictor of the change in walking. Figure 5 presents the change in walking by intervention status.

Figure 5. Change in walking by condition. Box: 1st-3rd quartile, whiskers: 1.5*interquartile range, circles: outliers.



Hypothesis 2

Our second hypothesis was supported: among intervention participants, the discrepancy between the individuals' prior beliefs and the presented outcomes was associated with their change in outcome expectancy. The coefficient for the interaction between intervention status and discrepancy score was -0.25 (SE $.07$, $t_{57} = -3.2$, $P < .01$), indicating that on average, after viewing the simulation, the beliefs of individuals with negative baseline beliefs became more positive while the beliefs

of those with overly optimistic baseline beliefs became more negative.

Hypotheses 3 and 4

Our secondary hypotheses were also supported: both groups increased behavioral intentions, mean difference 0.66 on a scale of 7 ($t_{62} = 4.5$, $P < .001$), and knowledge, mean difference 0.38 on a scale of 14 ($t_{62} = 2.4$, $P = .02$). Table 3 summarizes hypotheses 1 and 2 and their results. Table 4 summarizes hypotheses 3 and 4 and their results. Table 5 presents the means and standard deviations for all outcome measures.

Table 3. Summary of hypotheses and results.

Hypothesis	Model	Coefficient	SE	<i>t</i> value	<i>df</i>	<i>P</i> value
Walking will increase more in intervention participants	Linear model regressing postintervention walking on intervention status, preintervention walking adjusted for age and preintervention intent	61.0	30.5	1.9	58	.05
Among intervention participants, change in outcome expectancy will be a function of the discrepancy between prior beliefs and the presented outcome	Linear model regressing the change in outcome expectancy on an interaction term between intervention status and discrepancy score, adjusted for age and preintervention intent	-.25	.07	-3.213	57	<.01

Table 4. Summary of hypotheses and results

Hypothesis	Model	Mean difference	<i>t</i> value	<i>df</i>	<i>P</i> value
Both group will increase in behavioral intention	Paired <i>t</i> test comparing postintervention versus preintervention measure	0.66	4.5	62	<.001
Both groups will increase in diabetes-related knowledge	Paired <i>t</i> test comparing postintervention versus preintervention measure	0.38	2.4	62	.02

Table 5. Means (SD) for all outcome measures pre- and postintervention.

Outcome measure	Intervention status	Pre intervention	Postintervention
Walking (minutes)	Intervention	182.9 (245)	230.3 (262)
	Control	203.5 (203)	185.6 (193)
Outcome expectancy (scale 1–7)	Intervention	6.07 (1.1)	6.56 (.82)
	Control	6.37 (.89)	6.69(.55)
Behavioral intent (scale 1–7)	Intervention	4.79 (1.62)	5.62 (1.80)
	Control	5.53 (1.60)	6.03 (1.24)
Knowledge (scale 1–14)	Intervention	11.15 (2.3)	11.71 (2.14)
	Control	11.29 (1.95)	11.48 (2.18)

Qualitative Themes

We coded responses to qualitative questions into general themes and determined the proportion of each theme. When asked “What were the things that you liked about the simulation?” 31/65 of participants’ responses were coded as *informative*: these included comments such as “I thought the simulation was very clear” and “I think it was better than what I got in diabetes education.” Other themes that emerged were *surprise*: 11/65 participants commented that they were surprised at the effect of walking on the glucose curve. A third theme was *complications*: 7/65 participants reported liking the inclusion of pictures of individuals with complications; as 1 participant said, she felt that this was “important for people to see what might happen to them.” Finally, 5 participants reported that they had not seen or thought of glucose as a curve before, and 4 participants reported that they were previously unaware of the dawn phenomenon.

When asked “Were there things you did not like about the simulation?” most participants (52/65) answered “No.” Of those who provided specific negative feedback (13/65), 4 reported that the simulation contained “nothing new” or was “not interesting.” A total of 2 participants, both of whom worked

nights and slept during the day, reported feeling that the content of the simulation was not relevant to them. In addition, 3 reported not liking the music or voiceover, 1 reported not liking the glucose curves, 1 reported not liking the drawing task, 1 reported not liking the numeracy test, and 1 thought the simulation was too slow in the beginning.

When asked “Were there parts of the simulation you found confusing or that brought up questions in your mind?” most participants (59/65) answered “No.” Of those who provided specific feedback, 3 reported finding the drawing task confusing and 2 reported not understanding the meaning of the curves.

When asked “Are there things that were not in the simulation that you would like to see in a simulation like this?” 9 participants commented they would like to see the effect of different foods on the glucose curve, 5 wanted more information about how the disease progresses over time and whether it is reversible, 4 commented that they would like to see numbers on the curves, 3 commented that they would like to see more answers to the test questions addressed in the narrated simulation (not all the questions on the knowledge tests were addressed in the simulation), 2 commented that they would like to see the effect of insulin, and 2 control participants wanted to see the effect of exercise on the curve.

Although there was a small difference in the proportion of individuals who reported thinking about the glucose curves in the week following the simulation by condition (27/33 intervention participants, 22/32 controls), this difference was not significant ($\chi^2_1 = .88$, $P = .35$). When asked whether they thought about the glucose curves in the context of specific self-management behaviors, the proportions of all participants were as follows: when exercising (38/65), eating (35/65), and testing their blood sugar (30/65). There was no difference between groups in the incidence of thinking about the glucose curves in these contexts.

Discussion

This study had two main findings. First, intervention participants who completed an action plan for walking in the next week reported significantly more walking in the subsequent week than control participants. This finding is congruent with a large number of both laboratory and clinical studies that have found a positive impact of implementation intentions and action plans [12]. Our use of an action plan with simultaneous mental simulation of the plan is not novel. However, prior studies used a healthy university student population [23,24]; this study used an older diabetic population.

Our second main finding was that intervention participants' beliefs changed in accordance with the discrepancy between their prior beliefs and the outcomes presented in the simulation. The idea that computerized simulations could change outcome expectancies was suggested by Bandura in 1999 [35] and is in line with his earlier work demonstrating that individuals' beliefs change as a result of their observations of the effects of their own and others' behaviors [36]. We are unaware of any studies that have translated these ideas into a patient-facing intervention. We believe this finding suggests that computerized simulations could be used much more broadly to change individuals' health-related beliefs.

We are aware of only one other study involving glucose curves to promote physical activity among individuals with T2DM. Allen et al randomly assigned 52 individuals to one-on-one educational sessions [37]. The intervention session incorporated glucose curves to demonstrate the effect of physical activity on glucose. The session also included discussing the benefits of increased activity, assessing the individual's barriers to physical activity and self-efficacy for exercise, and providing an appropriate exercise prescription. The control session mentioned but did not stress physical activity as a self-management behavior for T2DM. At the 8-week follow-up, individuals in the intervention group had significantly greater improvements in self-efficacy for physical activity, accelerometer-measured physical activity, hemoglobin A1c, and body mass index. Both our study and Allen and colleagues' used glucose curves to promote physical activity in individuals with T2DM, but there are important differences. First, the proposed mechanisms are different: the Allen intervention was intended to increase physical activity by increasing participants' self-efficacy, while our intervention was intended to increase physical activity by changing outcome expectancies and implementing an action plan. Second, the degree of experimental control is different:

our study was a comparison between two computerized simulations that differed only in the inclusion of two components; Allen and colleagues' study compared in-person interventions that differed in many respects. We believe that these two studies, taken together, provide evidence that the outcome expectancies and self-efficacy of individuals with T2DM can be positively affected by modeling using glucose curves.

Implications for Translation

The results of this study highlight the potential for the translation of specific evidence from the psychology literature into the design of informatics-based behavioral interventions. We used an action planning intervention to facilitate subsequent action in intervention participants. This technique holds great promise to facilitate health-related behaviors, particularly in mobile phone-based interventions. In fact, recent evidence has shown that sending text message reminders of planned actions further facilitates the desired action [38]. We also presented potential outcomes as upward counterfactuals (how things might have been better) to maximize participants' behavioral intentions. This framing of information might be more widely used in consumer health informatics to increase user motivation; however, since we did not experimentally test this component of our intervention, further work is needed to test this idea.

Strengths

This study has several strengths. First, we employed prior findings in the psychological literature to design a brief, self-contained intervention and conducted a hypothesis-driven test of the efficacy of components of the intervention. Second, our use of glucose curves for both the presentation and elicitation of outcomes allowed for the measurement of individuals' outcome expectancies across three dimensions: the magnitude, duration, and direction of the effect. We believe this method is superior to the more common Likert scale measures of belief, and that a computer-based version of this drawing task could further improve upon the discrepancy score used in this study. A limitation of the discrepancy score used in this study is that it does not account for differences in the magnitude and duration of the individual's expectation (a larger discrepancy reflects a more inaccurate belief than a smaller discrepancy). A better measure of the discrepancy would be the difference in the area under the curve between the individual's curve and the presented outcome. This was not feasible using the complex curves drawn on paper in this study, but a computer-based version of the drawing task could easily calculate this difference.

Limitations

This study has limitations. First, our primary outcome measure, physical activity, was measured by self-report. Since all participants used the same measure, we do not believe this undermines the results; however, the true magnitude of the effect of our intervention on subsequent physical activity needs to be determined with objective measures in future work. Additionally, some of our participants did not represent the target population for this intervention: some participants possessed adequate diabetes-related numeracy, had positive outcome expectancies and intentions for exercise, were

knowledgeable about their disease, and were already physically active. We plan to address this issue in the future by integrating the intervention into diabetes education classes in target populations, particularly groups with newly diagnosed T2DM and low diabetes numeracy. The third limitation of this study was that the tests used to measure knowledge were not well aligned with the simulation's presentation of content. We developed the simulation around gists we considered important based on theory [39], our clinical experience, and pilot work. Available measures of diabetes-related knowledge, including those used in this study, measure an individual's knowledge of facts. Instruments measuring conceptual understanding of diabetes self-management are not available. In future work, simple simulations such as those used in this study could serve as a method to both teach and test understanding of diabetes-related concepts. A final limitation of this study is that, while we attempted to minimize the interaction between the investigator and the participant, some interaction was necessary (eg, the administration of the IPAQ). Further work is needed to determine the effectiveness of an entirely computer-based version of the intervention.

Future Research

The next generation of this intervention will test the effectiveness of personalizing the feedback provided in an interactive phone-based intervention. A phone-based intervention may facilitate integration of the simulation into the

user's daily life, may be easier to access than traditional diabetes education, which reaches a limited population [40], and might be less costly than an in-person intervention [41]. Recently, Fisher et al [42] and Polonsky et al [43] reported on an in-person intervention called the Structured Test Protocol. The core of this intervention was the estimation of the individual's glucose curves using 7-point glucose monitoring for 3 days. In their study, estimated curves facilitated shared decision making between patient and provider, resulting in a greater improvement in hemoglobin A1c, diabetes self-efficacy, autonomous motivation for diabetes care, and a more positive attitude toward self-monitoring of glucose than usual care [42]. This protocol concentrates the timing of self-monitoring but does not require a net increase in the volume of glucose monitoring [43], and therefore may be a cost-neutral and minimally invasive method to tailor the curves presented in the simulation. We hypothesize that the personalization of the presented curves, in combination with the personalization of the predicted effect of exercise (a subject of current research), may result in greater effectiveness of the intervention.

Conclusion

In this study we tested a simple form of a computer-based simulation. Participants' outcome expectancies changed in accordance with the discrepancy between their prior beliefs and the presented outcomes. In combination with action planning, the simulation positively affected short-term behavior.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Intervention simulation.

[[MOV File, 36MB - jmir_v14i3e71_app1.mov](#)]

Multimedia Appendix 2

Control simulation.

[[MOV File, 32MB - jmir_v14i3e71_app2.mov](#)]

Multimedia Appendix 3

Theory of Planned Behavior Questionnaire.

[[PDF File \(Adobe PDF File\), 51KB - jmir_v14i3e71_app3.pdf](#)]

References

1. American Diabetes Association. 2011. Diabetes Statistics URL: <http://www.diabetes.org/diabetes-basics/diabetes-statistics/> [accessed 2011-10-21] [[WebCite Cache ID 62bXVPKIV](#)]
2. Boulé NG, Haddad E, Kenny GP, Wells GA, Sigal RJ. Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials. JAMA 2001 Sep 12;286(10):1218-1227 [[FREE Full text](#)] [Medline: [11559268](#)]
3. Snowling NJ, Hopkins WG. Effects of different modes of exercise training on glucose control and risk factors for complications in type 2 diabetic patients: a meta-analysis. Diabetes Care 2006 Nov;29(11):2518-2527. [doi: [10.2337/dc06-1317](#)] [Medline: [17065697](#)]
4. Chudyk A, Petrella RJ. Effects of exercise on cardiovascular risk factors in type 2 diabetes: a meta-analysis. Diabetes Care 2011 May;34(5):1228-1237. [doi: [10.2337/dc10-1881](#)] [Medline: [21525503](#)]

5. Jakicic JM, Jaramillo SA, Balasubramanyam A, Bancroft B, Curtis JM, Mathews A, Look AHEAD Study Group. Effect of a lifestyle intervention on change in cardiorespiratory fitness in adults with type 2 diabetes: results from the Look AHEAD Study. *Int J Obes (Lond)* 2009 Mar;33(3):305-316. [doi: [10.1038/ijo.2008.280](https://doi.org/10.1038/ijo.2008.280)] [Medline: [19153582](https://pubmed.ncbi.nlm.nih.gov/19153582/)]
6. Nylen ES, Kokkinos P, Myers J, Faselis C. Prognostic effect of exercise capacity on mortality in older adults with diabetes mellitus. *J Am Geriatr Soc* 2010 Oct;58(10):1850-1854. [doi: [10.1111/j.1532-5415.2010.03068.x](https://doi.org/10.1111/j.1532-5415.2010.03068.x)] [Medline: [20929462](https://pubmed.ncbi.nlm.nih.gov/20929462/)]
7. Colberg S. Being active: a commentary. *Diabetes Educ* 2007;33(6):989-990. [doi: [10.1177/0145721707308479](https://doi.org/10.1177/0145721707308479)]
8. Zhao G, Ford ES, Li C, Balluz LS. Physical activity in U.S. older adults with diabetes mellitus: prevalence and correlates of meeting physical activity recommendations. *J Am Geriatr Soc* 2011 Jan;59(1):132-137. [doi: [10.1111/j.1532-5415.2010.03236.x](https://doi.org/10.1111/j.1532-5415.2010.03236.x)] [Medline: [21226683](https://pubmed.ncbi.nlm.nih.gov/21226683/)]
9. Lippa KD, Klein HA. Portraits of patient cognition: how patients understand diabetes self-care. *Can J Nurs Res* 2008 Sep;40(3):80-95. [Medline: [18947093](https://pubmed.ncbi.nlm.nih.gov/18947093/)]
10. Lippa KD, Klein HA, Shalin VL. Everyday expertise: cognitive demands in diabetes self management. *Hum Factors* 2008 Feb;50(1):112-120. [Medline: [18354975](https://pubmed.ncbi.nlm.nih.gov/18354975/)]
11. Klein H, Lippa K. Type 2 diabetes self-management: controlling a dynamic system. *J Cogn Eng Decis Making* 2008;2(1):48-62. [doi: [10.1518/155534308X284363](https://doi.org/10.1518/155534308X284363)]
12. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement: a meta-analysis of effects and processes. *Adv Exp Soc Psychol* 2006;38:69-119. [doi: [10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)]
13. Hedbrant J, Ludvigsson J. Use of computer simulator training in the education of diabetic teenagers. *Pract Diabetes Int* 1995;12(1):18-21. [doi: [10.1002/pdi.1960120109](https://doi.org/10.1002/pdi.1960120109)]
14. Tatti P, Lehmann ED. A prospective randomised-controlled pilot study for evaluating the teaching utility of interactive educational diabetes simulators. *Diabetes Nutr Metab* 2003 Feb;16(1):7-23. [Medline: [12848301](https://pubmed.ncbi.nlm.nih.gov/12848301/)]
15. Fritschi C, Quinn L, Penckofer S, Surdyk PM. Continuous glucose monitoring: the experience of women with type 2 diabetes. *Diabetes Educ* 2010;36(2):250-257. [doi: [10.1177/0145721709355835](https://doi.org/10.1177/0145721709355835)] [Medline: [20016057](https://pubmed.ncbi.nlm.nih.gov/20016057/)]
16. Plotnikoff RC, Lippke S, Courneya K, Birkett N, Sigal R. Physical activity and diabetes: an application of the theory of planned behaviour to explain physical activity for Type 1 and Type 2 diabetes in an adult population sample. *Psychol Health* 2010 Jan;25(1):7-23. [doi: [10.1080/08870440802160984](https://doi.org/10.1080/08870440802160984)] [Medline: [20391204](https://pubmed.ncbi.nlm.nih.gov/20391204/)]
17. Plotnikoff R. Physical activity and social cognitive theory: a test in a population sample of adults with type 1 or type 2 diabetes. *Appl Psychol* 2008;57(4):628-643. [doi: [10.1111/j.1464-0597.2008.00344.x](https://doi.org/10.1111/j.1464-0597.2008.00344.x)]
18. Glasgow RE, Hampson SE, Strycker LA, Ruggiero L. Personal-model beliefs and social-environmental barriers related to diabetes self-management. *Diabetes Care* 1997 Apr;20(4):556-561. [Medline: [9096980](https://pubmed.ncbi.nlm.nih.gov/9096980/)]
19. Broadbent E, Donkin L, Stroh JC. Illness and treatment perceptions are associated with adherence to medications, diet, and exercise in diabetic patients. *Diabetes Care* 2011 Feb;34(2):338-340. [doi: [10.2337/dc10-1779](https://doi.org/10.2337/dc10-1779)] [Medline: [21270191](https://pubmed.ncbi.nlm.nih.gov/21270191/)]
20. Williams DM, Anderson ES, Winett RA. A review of the outcome expectancy construct in physical activity research. *Ann Behav Med* 2005 Feb;29(1):70-79. [doi: [10.1207/s15324796abm2901_10](https://doi.org/10.1207/s15324796abm2901_10)] [Medline: [15677303](https://pubmed.ncbi.nlm.nih.gov/15677303/)]
21. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychol Bull* 2006 Mar;132(2):249-268. [doi: [10.1037/0033-2909.132.2.249](https://doi.org/10.1037/0033-2909.132.2.249)] [Medline: [16536643](https://pubmed.ncbi.nlm.nih.gov/16536643/)]
22. Webb T, Sheeran P. How do implementation intentions promote goal attainment? A test of component processes. *J Exp Soc Psychol* 2007;43(2):295-302. [doi: [10.1016/j.jesp.2006.02.001](https://doi.org/10.1016/j.jesp.2006.02.001)]
23. Knauper B, Roseman M, Johnson P, Krantz L. Using mental imagery to enhance the effectiveness of implementation intentions. *Curr Psychol* 2009;28:181-186.
24. Knäuper B, McCollam A, Rosen-Brown A, Lacaille J, Kelso E, Roseman M. Fruitful plans: adding targeted mental imagery to implementation intentions increases fruit consumption. *Psychol Health* 2011 May;26(5):601-617. [doi: [10.1080/08870441003703218](https://doi.org/10.1080/08870441003703218)] [Medline: [21337259](https://pubmed.ncbi.nlm.nih.gov/21337259/)]
25. Gibson B, Weir C. Development and preliminary evaluation of a simulation-based diabetes education module. *AMIA Annu Symp Proc* 2010;2010:246-250. [Medline: [21346978](https://pubmed.ncbi.nlm.nih.gov/21346978/)]
26. Markman KD, McMullen MN. A reflection and evaluation model of comparative thinking. *Pers Soc Psychol Rev* 2003;7(3):244-267. [Medline: [12788690](https://pubmed.ncbi.nlm.nih.gov/12788690/)]
27. Smallman R, Roese NJ. Counterfactual Thinking Facilitates Behavioral Intentions. *J Exp Soc Psychol* 2009 Jul;45(4):845-852. [doi: [10.1016/j.jesp.2009.03.002](https://doi.org/10.1016/j.jesp.2009.03.002)] [Medline: [20161221](https://pubmed.ncbi.nlm.nih.gov/20161221/)]
28. Sheeran P, Webb TL, Gollwitzer PM. The interplay between goal intentions and implementation intentions. *Pers Soc Psychol Bull* 2005 Jan;31(1):87-98. [doi: [10.1177/0146167204271308](https://doi.org/10.1177/0146167204271308)] [Medline: [15574664](https://pubmed.ncbi.nlm.nih.gov/15574664/)]
29. Huizinga MM, Elasy TA, Wallston KA, Cavanaugh K, Davis D, Gregory RP, et al. Development and validation of the Diabetes Numeracy Test (DNT). *BMC Health Serv Res* 2008;8:96 [FREE Full text] [doi: [10.1186/1472-6963-8-96](https://doi.org/10.1186/1472-6963-8-96)] [Medline: [18452617](https://pubmed.ncbi.nlm.nih.gov/18452617/)]
30. Ajzen I. University of Massachusetts Amherst. Constructing a Theory of Planned Behavior Questionnaire URL: <http://people.umass.edu/ajzen/pdf/tpb.measurement.pdf> [accessed 2012-04-23] [WebCite Cache ID 6798bwmnz]
31. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr* 2006 Sep;9(6):755-762. [Medline: [16925881](https://pubmed.ncbi.nlm.nih.gov/16925881/)]

32. Fitzgerald JT, Funnell MM, Hess GE, Barr PA, Anderson RM, Hiss RG, et al. The reliability and validity of a brief diabetes knowledge test. *Diabetes Care* 1998 May;21(5):706-710. [Medline: [9589228](#)]
33. Berikai P, Meyer PM, Kazlauskaitė R, Savoy B, Kozik K, Fogelfeld L. Gain in patients' knowledge of diabetes management targets is associated with better glycemic control. *Diabetes Care* 2007 Jun;30(6):1587-1589. [doi: [10.2337/dc06-2026](#)] [Medline: [17372160](#)]
34. R Foundation for Statistical Computing. The Comprehensive R Archive Network URL: <http://cran.r-project.org/> [accessed 2012-04-23] [WebCite Cache ID 6798f0Utx]
35. Bandura A. Social cognitive theory: an agentic perspective. *Asian J Soc Psychol* 1999;2:21-41. [doi: [10.1111/1467-839X.00024](#)]
36. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977 Mar;84(2):191-215. [Medline: [847061](#)]
37. Allen NA, Fain JA, Braun B, Chipkin SR. Continuous glucose monitoring counseling improves physical activity behaviors of individuals with type 2 diabetes: A randomized clinical trial. *Diabetes Res Clin Pract* 2008 Jun;80(3):371-379. [doi: [10.1016/j.diabres.2008.01.006](#)] [Medline: [18304674](#)]
38. Prestwich A, Perugini M, Hurling R. Can implementation intentions and text messages promote brisk walking? A randomized trial. *Health Psychol* 2010 Jan;29(1):40-49. [doi: [10.1037/a0016993](#)] [Medline: [20063934](#)]
39. Reyna VF. A theory of medical decision making and health: fuzzy trace theory. *Med Decis Making* 2008;28(6):850-865. [doi: [10.1177/0272989X08327066](#)] [Medline: [19015287](#)]
40. Centers for Disease Control and Prevention. 2011 Nov 14. Data 2010: The Healthy People 2010 Database URL: <http://wonder.cdc.gov/data2010/focus.htm> [accessed 2012-01-29] [WebCite Cache ID 654DQ6MDx]
41. Tate DF, Finkelstein EA, Khavjou O, Gustafson A. Cost effectiveness of internet interventions: review and recommendations. *Ann Behav Med* 2009 Aug;38(1):40-45. [doi: [10.1007/s12160-009-9131-6](#)] [Medline: [19834778](#)]
42. Fisher L, Polonsky W, Parkin C, Jelsofsky Z, Wagner R. Structured SMBG promotes positive changes in self-management attitudes in non-insulin treated T2DM: STeP Study Results. 2011 Presented at: 71st Scientific Sessions of the American Diabetes Association; June 24-28, 2011; San Diego, CA, USA.
43. Polonsky WH, Fisher L, Schikman CH, Hinnen DA, Parkin CG, Jelsofsky Z, et al. Structured self-monitoring of blood glucose significantly reduces A1C levels in poorly controlled, noninsulin-treated type 2 diabetes: results from the Structured Testing Program study. *Diabetes Care* 2011 Feb;34(2):262-267. [doi: [10.2337/dc10-1732](#)] [Medline: [21270183](#)]

Abbreviations

IPAQ: International Physical Activity Questionnaire

T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 15.10.11; peer-reviewed by R Mceachan, H Witteman; comments to author 27.11.11; revised version received 30.01.12; accepted 22.02.12; published 10.05.12.

Please cite as:

Gibson B, Marcus RL, Staggers N, Jones J, Samore M, Weir C

Efficacy of a Computerized Simulation in Promoting Walking in Individuals With Diabetes

J Med Internet Res 2012;14(3):e71

URL: <http://www.jmir.org/2012/3/e71/>

doi: [10.2196/jmir.1965](#)

PMID: [22576226](#)

©Bryan Gibson, Robin L Marcus, Nancy Staggers, Jason Jones, Matthew Samore, Charlene Weir. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org/>), 10.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Short-term Effects of a Smoking Prevention Website in American Indian Youth

Deborah J Bowen¹, PhD; Patricia Nez Henderson², PhD; Jessica Harvill³, MPH; Dedra Buchwald⁴, MD

¹School of Public Health, Boston University, Boston, MA, United States

²Black Hills Center for American Indian Health, Rapid City, SD, United States

³State of Alaska Department of Health and Social Services, Section of Chronic Disease, Anchorage, AK, United States

⁴University of Washington School of Medicine, Partnerships for Native Health, Seattle, WA, United States

Corresponding Author:

Deborah J Bowen, PhD
School of Public Health
Boston University
801 Massachusetts Avenue
Boston, MA, 02118
United States
Phone: 1 617 638 5205
Fax: 1 617 638 4483
Email: dbowen@bu.edu

Abstract

Background: The rate of smoking commercial tobacco products among American Indian youth is double the rate for white youth. Interventions are needed to reduce this disparity.

Objective: To test the feasibility of a Web-based intervention to influence attitudes toward and intentions about smoking cigarettes among American Indian youth who attended a Native summer camp in the Northern Plains.

Methods: The study website, the SmokingZine, was originally developed and tested in Canadian youth, then adapted to be appropriate for American Indian youth. We conducted a randomized controlled trial to test the influence of exposure to the adapted SmokingZine website on smoking attitudes and behaviors among American Indian youth 12–18 years of age. Participants assigned to the intervention group were given access to the website for 1 hour per day during their camp experience and asked to sign in to the site and use it. Control group participants were not given access to the site.

Results: A total of 52% of intervention youth signed in to the website at least once. Among nonsmokers, intentions to try a cigarette in the intervention group declined from 16% to 0%, and increased from 8% to 25% in the control group ($P < .05$). Compared with the control group, youth in the intervention group were more likely to help others quit (21 percentage point change in intervention versus no change in control; $P < .05$) and had less positive attitudes about the drug effects of smoking (–0.19 change in intervention versus 0.67 in control; $P < .05$).

Conclusion: These data indicate that SmokingZine needs more long-term, rigorous investigation as a way to keep American Indian youth from becoming regular smokers. Because the intervention group could use computers only 1 hour per day, increasing access might result in more visits and a greater effect of the website on smoking behaviors.

(*J Med Internet Res* 2012;14(3):e81) doi:[10.2196/jmir.1682](https://doi.org/10.2196/jmir.1682)

KEYWORDS

Smoking prevention; Native American; eHealth; intervention; cultural appropriateness

Introduction

Smoking tobacco is responsible for more premature morbidity and mortality in the United States than any other behavioral risk factor [1]. Nationally, rates of smoking, as well as other forms of tobacco use, such as chewing tobacco, are substantially

higher among American Indians than white people [2,3] and have not followed the downward trend observed over the past decade in the general US population [4,5]. Of particular concern, a recent report from the US Surgeon General on tobacco use among racial and ethnic minority people in the United States indicated that smoking rates among American Indian youth are

double that of their white counterparts [4] and remain high even compared with their peers in other racial and ethnic minority groups [4].

High rates of smoking among American Indian youth may be due, in part, to the sovereign status of federally recognized tribes. Because tribal lands are often not subject to state tobacco control policies, American Indian youth have access to tobacco products at a very early age. Early initiation often leads to addiction and long-term abuse of tobacco [6,7], which in turn contribute to cardiovascular disease and lung cancer, the leading causes of death in the American Indian population [6,7]. Although these worrisome data highlight the need to develop or identify successful programs to both prevent and reduce smoking among American Indian youth, a Medline search conducted before initiating this project produced no adequately evaluated programs that targeted Native youth.

Increasing attention has focused on using the Web for smoking cessation support, including delivery of information, discussion groups, cognitive behavioral treatment, and self-help materials [8-14]. The SmokingZine website was originally developed explicitly for testing in urban Canadian youth [15]. Since 1995, the TeenNet project [14,16] at the University of Toronto has pioneered a combination of website development, community mobilization, and action research involving young people from diverse backgrounds in all stages of program design, development, and dissemination. In 2002, TeenNet began a randomized controlled trial of SmokingZine involving more than 1400 smoking and nonsmoking adolescents from participating Toronto-area high schools [8]. SmokingZine was both the largest randomized trial of an eHealth intervention for behavior change and one of the largest studies on adolescent smoking cessation [8,11]. In another clinical trial, the same intervention was efficacious in reducing smoking among middle-school youth in public schools [15].

Because of the success of the SmokingZine trial with youth from diverse backgrounds, adapting and testing this program for Native youth seemed promising. Because young people use the Web for health information [10], we conducted a randomized controlled study of this Web-based youth smoking prevention and cessation resource, SmokingZine [8,11], among American Indians 12–18 years of age attending summer camp in South Dakota. The outcomes of interest to our study were short-term smoking behaviors, attitudes about smoking, and usability of the Web tool [12]. Our findings are intended to identify issues for a full-scale efficacy study of the SmokingZine intervention in a larger sample.

Methods

Overview

We conducted a small randomized feasibility trial in one geographical site to test the 1-month effects of a Web-based intervention on smoking behaviors among American Indian youth 12–18 years of age. All participants completed a baseline survey and were then randomly assigned to one of two conditions: intervention (50%) or control. Intervention participants were asked to sign in to the website and use it, and

were provided with regular access to the site during their camp experience. Control participants were not given access to the site. All participants completed a follow-up survey 1 month after randomization. We assessed and compared smoking behavior in both groups. In addition, we assessed attitudes and social behaviors that could be expected to change as a result of the intervention exposure. All control participants received access to the intervention after completion of the study. This study is reported in line with CONSORT requirements, and a flow chart and checklist were included in the review of the manuscript. We did not register this trial prospectively, as it was a feasibility trial, not a fully powered randomized trial.

Setting and Participants

Each year more than 150 American Indian 6th to 12th graders from across South Dakota are selected to participate in a 6-week residential summer enrichment program in Rapid City, South Dakota. All nine Indian reservations located in South Dakota participate in the program, whose goal is to prepare students for success in college. During the 6 weeks, students attend classes in mathematics, science, English, computer use, and life skills. All camp attendees were eligible for this study, with a target sample size of 100 youth, based on a sample size calculation with power of 80% and a moderate effect size in intentions to smoke in the future.

Recruitment and Consent Procedures

Working closely with the director of the summer program, we arranged to insert an informational sheet about the study into all correspondence mailed to the 164 selected students before the program started. During registration, our research staff set up an informational booth in the lobby of the students' dormitory to describe the program to students and their parents. For students interested in participating, a consent form was read to both parents and students. A total of 113 students provided consent, completed the baseline assessment described below, and were randomly assigned to the intervention or control arm. Randomization was completed by assigning each consented and assessed youth to a study identification number that had been preassigned to either intervention or control status. Field staff were unaware of the preassignment.

Intervention

Although it has a small cessation focus, the Web-based SmokingZine is primarily a smoking prevention intervention. It consists of a series of smoking educational modules that include culturally relevant contexts (a summer powwow) and images relevant to Native youth (eg, eagles, feathers, drums, and outdoor scenes). Our methods for developing and culturally adapting SmokingZine for use in this study have been previously published [16]. Briefly, we conducted talking circles with Native youth to ascertain their knowledge and attitudes about tobacco use, their opinions of SmokingZine, and their recommendations for changes. Talking circle participants provided useful information on the differences between ceremonial and commercial tobacco use, as well as their personal patterns and motivations regarding tobacco use. They also offered guidance on changes to the site that would make it more fun to use and more culturally appropriate. Examples of suggestions are

replacing urban images with images of outdoor settings familiar to diverse Native youth, and replacing images of a shopping mall with images of a powwow. The research team made changes based on participants' suggestions.

After randomization, control group participants were told that they would receive information about smoking at the end of the 6-week program. Intervention youth were given usernames and passwords to access the SmokingZine site and were instructed to use it as much as they wanted. Our intent was for the intervention participants to visit the site multiple times, with each visit lasting 10–20 minutes. The use of unique usernames and passwords enabled the research team to broadly track usage patterns of each participant over time. The research assistant maintained a list that linked each username to an identification number printed on the study questionnaire, enabling us to link site usage to baseline and follow-up data.

Each day of the 6-week program, students had 1 hour of computer time. Those in the intervention were encouraged to visit the SmokingZine site during this time. On the initial visit, the research assistant helped participants navigate to the site. Once participants signed in to the site, they were asked about their current smoking status. Then, based on their response, they were guided to educational modules about smoking prevention or cessation. The underlying message of these modules reinforced the values of nonsmokers and created a set of conditions to enhance smokers' motivation and commitment to change. Toward the end of each session, modules helped smokers identify goals for behavior change and explore barriers to change, as well as providing advice and strategies where appropriate. During subsequent sessions, participants received additional information about smoking prevention or cessation.

Assessments

All intervention and control youth were asked to complete a 30-minute paper-and-pencil survey on current smoking habits, intentions to smoke, and attitudes about smoking, both immediately after randomization and again 1 month later. We also approached youth who did not complete the entire survey to provide us with smoking outcomes only, to maximize response to a minimal outcome dataset.

The 87-item survey, called A Smoking Prevention Interactive Experience (ASPIRE) instrument, consisted of sociodemographic items, several scales, and individual items. Many of the items were derived from existing instruments, and others were created specifically for an interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high-school students [9]. Attitudes were measured by a series of questions asking for the opinions of youth on various smoking-related issues. The primary outcome of this feasibility study was short-term smoking behaviors; secondary outcomes were attitudes about smoking and intentions to smoke.

At baseline, on the basis of questionnaire answers, students were categorized as current smokers (smoke every other week, smoke less than a pack a week, smoke a pack a week, smoke more than a pack a week, smoke a pack a day, or smoke more than a pack a day), former smokers (used to smoke regularly but quit), or nonsmokers, including never-smokers (never

smoked even part of a cigarette) and experimenters (only smoked part of a cigarette or smoked only a few times).

Analyses

We compared baseline demographic data between intervention and control participants, using *t* tests for continuous variables and chi-square tests for proportions to confirm that randomization was accomplished. We also examined website usage data for intervention participants to assess patterns and frequency of use.

Next, we examined changes in survey responses from baseline to 1-month follow-up in key smoking and related outcomes. We focused these analyses on selected intermediate outcomes of long-term smoking prevention, such as short-term smoking behavior changes, intentions to smoke in the future, and social actions that participants could take regarding smoking. We compared each of these variables from baseline to follow-up between intervention and control participants. Relatively few participants were missing at follow-up and, therefore, we used no imputation method.

We also performed a factor analysis on the attitude questions in the ASPIRE instrument to determine whether we could identify any groupings of questions for analysis. We identified three clusters, based on factor loadings of greater than 0.5, independent of loading on other factors. These were social issues (3 items, eg, "Kids who smoke have more friends"), drug effects (3 items, eg, "Smoking cigarettes relieves tension"), and negative effects (9 items, eg, "Cigarette smoking is addictive"). We averaged the items in each cluster to form a score for each area. We then compared responses to attitudinal scores between intervention and control youth from baseline to follow-up.

Results

The study team invited 164 young people to participate, and recruited and randomized 113 (68.9%). The mean age of participants in the intervention and control arms did not differ (mean of 14.8 vs 14.4 years). Of the 113 who completed the baseline measures, we were able to collect smoking follow-up data on 102 (90.3%). The intervention and control groups did not differ in the proportion who provided follow-up data ($n = 49$, 92% versus $n = 44$, 90%). Likewise, baseline demographic data and smoking behavior were similar between youth who completed and those who did not complete the follow-up assessment.

In terms of use, 10 (52%) intervention youth signed in to the website at least once. Of these, 39 (80%) used the site only once, and 5 (20%) used it 2 or more times. Overall, intervention users rated the site positively on the survey: 16 (33%) rated it as very useful and 27 (54%) rated it as somewhat useful. A large proportion of intervention users indicated that use of the site made them think very differently 22 (44%) or somewhat differently 19 (38%) about smoking. Finally, most of the intervention users indicated that use of the site would very much (36, 74%) or somewhat (3, 11%) keep them from smoking in the future.

Table 1 presents the characteristics of the 113 participants randomly assigned in the study. Baseline demographic characteristics did not differ across the intervention and control groups. Most participants reported their ethnic status as Native (95/113, 84%), with a preponderance of girls over boys (69, 61%), and a majority with at least one smoking parent (75, 66%).

Table 1. Characteristics of American Indian youth participating in a feasibility trial of the SmokingZine website.^a

Characteristic	Intervention group (n = 64), n (%)	Control group (n = 49), n (%)	P value
Female	34 (53%)	35 (71%)	>.22
Race			>.86
American Indian	58 (91%)	37 (76%)	
White	2 (3%)	7 (14%)	
Other	4 (6%)	5 (10%)	
Age (years)			>.35
£13	7 (11%)	6 (21%)	
14–26	57 (89%)	43 (88%)	
≥27	0	0	
Household size			>.31
£5	31 (48%)	28 (57%)	
≥6	33 (51%)	21 (43%)	
Mother smokes			>.18
Yes	28 (45%)	30 (61%)	
Never smoked	16 (26%)	13 (27%)	
Quit smoking	16 (26%)	6 (12%)	
Don't know	2 (3%)	0	
Father smokes			>.40
Yes	25 (40%)	16 (33%)	
Never smoked	16 (25%)	14 (29%)	
Quit smoking	15 (24%)	10 (21%)	
Don't know	7 (11%)	7 (15%)	
Parental marital status			>.71
Married	27 (42%)	22 (45%)	
Separated	15 (23%)	12 (25%)	
Divorced	12 (19%)	7 (14%)	
Single	7 (11%)	5 (10%)	
Widowed	1 (2%)	1 (2%)	
One or both deceased	2 (3%)	2 (4%)	

^a Percentages are computed with 1–2 missing values excluded; sum of percentages may not total 100% due to rounding.

Table 2 presents data on smoking status and intent for the intervention and control groups at baseline and 1-month follow-up. Differences including smokers and former smokers were not tested statistically due to small sample sizes. Rates of current smoking at baseline among all youth were relatively low: almost 80% were never-smokers. The proportions of youth in the *never*, *current*, and *former* categories of smoking did not change from baseline to follow-up. The number of intervention

youth who reported helping someone quit smoking increased over time. In contrast, the number in the control group did not change (6 versus 0, $P < .05$). Likewise, among nonsmokers, the number who intended to try a cigarette in the intervention group declined from 7 to 0 and increased from 3 to 8 in the control group ($P < .05$). Finally, the other smoking-related variables did not differ across the intervention and control groups.

Table 2. Effects of exposure to the SmokingZine intervention on smoking-related variables.^a

Question	Intervention group		Control group		P value
	Baseline (n = 64), n (%)	Follow-up (n = 59), n (%)	Baseline (n = 49), n (%)	Follow-up (n = 44), n (%)	
Smoking status					.58
Nonsmoker	51 (85%)	47 (83%)	38 (83%)	35 (80%)	
Current smoker	5 (8%)	3 (5%)	2 (4%)	2 (4%)	
Former smoker	4 (6%)	7 (12%)	6 (13%)	6 (15%)	
Missing data	4	2	3	2	
All youth ...Tried to help someone quit smoking					<.01
Yes	46 (70%)	52 (91%)	28 (58%)	28 (58%)	
No	20 (30%)	5 (9%)	20 (42%)	15 (42%)	
Former smokers ...At any time during the next year, do you think you will smoke?					.62
Yes	1 (10%)	2 (29%)	6 (100%)	6 (100%)	
No	9 (90%)	5 (71%)	0	0	
Nonsmokers ...In the next year, do you think you will try smoking a cigarette?					<.02
Yes	7 (16%)	0	3 (8%)	8 (25%)	
No	37 (84%)	38 (100%)	34 (92%)	27 (75%)	
Never-smokers ...If one of your best friends offered you a cigarette, would you smoke it?					.26
Yes	1 (2%)	2 (5%)	1 (3%)	0	
No	43 (98%)	42 (96%)	36 (97%)	32 (100%)	

^a Percentages were computed with 1–3 missing values excluded; sum may not total 100% due to rounding.

Table 3 presents data on the attitudinal changes from baseline to follow-up in both the intervention and control groups. Scores for the drug effects and negative effects clusters (described above), but not for social issues, changed significantly from baseline to follow-up in the intervention group but not in the

control group, in a direction consistent with the intended effects. Intervention youth felt less positively than the control youth about the drug effects and more negatively about the adverse effects of tobacco.

Table 3. Key short-term attitudes (range 1–4) related to smoking for the intervention and control groups.^a

Attitude cluster	Intervention group (n = 64)		Comparison group (n = 44)		P value
	Mean	SD	Mean	SD	
Social issues					.23
Baseline	1.32	0.18	1.20	0.26	
Follow-up	1.11	0.27	1.31	0.28	
Drug effects					<.04
Baseline	1.86	1.1	1.67	1.1	
Follow-up	1.1	1.2	1.77	1.2	
Negative effects					<.02
Baseline	4.20	1.7	4.05	1.9	
Follow-up	4.8	1.4	4.01	1.8	

^a Means are computed with all follow-up missing values excluded; n may not total 100% due to rounding or missing variables.

Discussion

The purpose of the present study was to test several aspects of the feasibility of the adapted Web-based SmokingZine

intervention for American Indian youth. The intervention was feasible and acceptable for at least half of the participants, and at least somewhat effective for American Indians 12–18 years of age. The trial was successfully implemented with little

oversight or structure, among staff and youth generally unfamiliar with research projects or randomized designs. Moreover, most students (69%) in the summer camp participated in the study, indicating that a larger randomized test of this, or a similar, Web-based intervention is logistically possible. This evidence of feasibility is critical, because concerns about the adequacy of prior study designs have plagued the literature on youth smoking prevention [13]. We also were pleased that a substantial proportion of the intervention group found the website interesting and useful.

Only 52% used the site, although a much smaller proportion used it on more than one occasion. This is a common occurrence among Web-based studies of youth, where actual use of the website in randomized trials is often low [17]. In this study the investigators identified some engagement strategies for youth to increase their use of and engagement in the Web-based activities, and this is a promising avenue for future research and implementation. Future studies should consider strategies to increase the proportion of randomly assigned participants who visit the SmokingZine website more than once, for instance by structuring or scheduling use, providing incentives for use, or increasing access to computers. Qualitative work has identified the multiple ways in which youth engage in using eHealth technologies to gain information [18], and some of these ways are relevant to the present application. For example, finding personalized and tailored information was a need expressed by youth in the focus groups, and the SmokingZine website could be more tailored to individual users' needs and data and marketed as such.

In regard to marketing, in our study, the moderate amount of website use might have resulted from the strict limitation on computer access, suggesting that liberalizing the amount of computer time might enhance the effect of the intervention. If the program were implemented in schools, there would be supervised and required time for sessions on the computer that would likely promote the visiting of the website by youth during school hours.

Given our experiences with this study, we propose that a group design would be ideal, given the opportunity for sharing site passwords as well as information and materials pertinent to the intervention among youth in the same physical location. The popularity among youth of social networking sites such as Facebook also suggests that a group design would be appropriate for a larger trial. As well, in a study with a longer follow-up period, preexisting friendships and social networks would result in youth using the site together. We did not observe joint use in our study, but in settings where computers are publicly available, such as schools and community centers, joint use could occur. The use of group randomized trials, in which the unit of intervention is the group, would help to minimize this problem. Finally, biochemical validation could be used to verify smoking outcomes in any future rigorous efficacy trials aimed at American Indian youth.

In the short run, the intervention did not directly affect smoking behavior, although it did alter intentions to use tobacco among never-smokers. This finding is encouraging for future efficacy trials, as the relatively high rates of smoking in later adolescence

occur in both early- and later-starting smokers. This type of intervention may thus be best suited for young people who are still irregular smokers or nonsmokers at the time of exposure. The intervention did change social interactions involving smoking, such as offering to help others quit. This may be the most intriguing and important finding, as tobacco use is a socially driven phenomenon [19] that could be prevented by breaking down socially normative perspectives on tobacco.

Given our small sample size, the lack of effects on smokers may simply indicate lack of power coupled with low baseline rates of smoking. These low baseline rates, in turn, likely reflect the biases related to being selected to attend a camp for American Indian youth. For example, the students had to have a high grade-point average and were highly recommended by their school counselor and teachers. Smoking is more frequent among Native youth with lower academic performance, who perhaps were not widely represented at this camp [16]. In other community settings, we would be more likely to encounter a wider variety of Native youth, including more regular smokers. This is a major limitation to the generalizability of the present study data, in that less achievement-oriented Native youth might not respond to this website. The structured nature of the summer camp is also a limitation of the present study, as youth in the camp are not near their regular and ongoing social influences, and so the artificial structure of the camp interactions might alter the effects of the intervention.

Besides the highly selected sample of American Indian youth attending summer camp, other limitations of this feasibility study include the short duration of the follow-up and the lack of qualitative data to enhance our understanding of quantitative survey findings. As well, a key limitation was the camp's rigorous control of access to the Web, allowing youth only 1 hour of computer use each day. We expect that participants neither visited the site as much nor stayed as long as they might have with longer or unlimited Web access, thus diluting the potential impact of the intervention. There was always the potential in this individually randomized design for contamination to occur between intervention and comparison youth, discussing smoking and even using the website together. Finally, despite our best efforts, the randomization intended to be evenly divided between intervention and comparison youth was not, and so more rigorous randomization procedures will need to be instituted for the next study. Even so, strengths include the use of a previously tested intervention, a documented adaptation process, and a rigorous feasibility design.

Future efforts should pay closer attention to cessation among subsets of youth smokers, because American Indians who smoke in youth are most likely to become regular adult smokers [20]. If cessation needs more emphasis, the intervention website will need to be enhanced.

These findings partially support the potential of the SmokingZine tool for future research. We believe this feasibility study has set the stage for future intervention research into tobacco use reduction among American Indian and Alaska Native youth. Online tobacco control efforts have started to emerge [21-24], but none to date have been designed for or targeted at American Indians and Alaska Natives. In a similar

vein, some authors have called for more use of eHealth and outside the Indian Health Service [19]. technology to enhance the health of Native people, both inside

Acknowledgments

We are deeply grateful for the collegiality of Dr Harvey Skinner and his team at the University of Toronto. With this generous help, we have been able to adapt their intervention for our population.

This work was supported by the National Cancer Institute 1U01 CA114642 to D.S.B.

Conflicts of Interest

None declared.

References

1. Anonymous. The 2004 United States Surgeon General's Report: The Health Consequences of Smoking. N S W Public Health Bull 2004;15(5-6):107. [Medline: [15543245](#)]
2. Centers for Disease Control and Prevention (CDC). Cigarette smoking among adults--United States, 1997. MMWR Morb Mortal Wkly Rep 1999 Nov 5;48(43):993-996 [FREE Full text] [Medline: [10577493](#)]
3. Centers for Disease Control and Prevention. 2000. American Indians and Alaska Natives and Tobacco URL: http://www.in.gov/isdh/tpc/files/research_66.pdf [accessed 2012-05-28] [WebCite Cache ID 67zxOiimd]
4. US Department of Health and Human Services. 1998. Tobacco Use Among US Racial/Ethnic Minority Groups: African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, Hispanics: A Report of the Surgeon General URL: <http://profiles.nlm.nih.gov/ps/access/NNBBFQ.pdf>
5. Warner KE. Smoking and health: a 25-year perspective. Am J Public Health 1989 Feb;79(2):141-143. [Medline: [2913830](#)]
6. National Center for Health Statistics. 2000. Health, United States, 2000 URL: <http://www.cdc.gov/nchs/data/hsus/hsus00.pdf> [accessed 2012-05-28] [WebCite Cache ID 67zxTUXsl]
7. US Department of Health and Human Services. 1988. The Health Consequences of Smoking: Nicotine Addiction: A Report of the Surgeon General URL: <http://profiles.nlm.nih.gov/ps/access/NNBBZD.pdf>
8. Feil EG, Noell J, Lichtenstein E, Boles SM, McKay HG. Evaluation of an Internet-based smoking cessation program: lessons learned from a pilot study. Nicotine Tob Res 2003 Apr;5(2):189-194. [Medline: [12745491](#)]
9. Gray NJ, Klein JD. Adolescents and the internet: health and sexuality information. Curr Opin Obstet Gynecol 2006 Oct;18(5):519-524. [doi: [10.1097/01.gco.0000242954.32867.76](#)] [Medline: [16932046](#)]
10. Lenert L, Muñoz RF, Stoddard J, Delucchi K, Bansod A, Skoczen S, et al. Design and pilot evaluation of an internet smoking cessation program. J Am Med Inform Assoc 2003;10(1):16-20 [FREE Full text] [Medline: [12509354](#)]
11. Ribisl KM, Kim AE, Williams RS. Web sites selling cigarettes: how many are there in the USA and what are their sales practices? Tob Control 2001 Dec;10(4):352-359 [FREE Full text] [Medline: [11740027](#)]
12. Skinner H, Morrison M, Bercovitz K, Haans D, Jennings MJ, Magdenko L, et al. Using the Internet to engage youth in health promotion. Promot Educ 1997 Dec;4(4):23-25. [Medline: [9560853](#)]
13. Tauli M, Bush N, Bowen DJ, Forquera R. Adaptation of a smoking cessation and prevention website for urban American Indian/Alaska Native youth. J Cancer Educ 2010 Mar;25(1):23-31. [doi: [10.1007/s13187-009-0004-2](#)] [Medline: [20082169](#)]
14. Program Training and Consultation Centre. 2009. The Smoking Zine: Using the Internet for Smoking Prevention With Youth URL: <https://www.ptcc-cfc.on.ca/cms/One.aspx?portalId=97833&pageId=114647>
15. Mentor International. The Smoking Zine URL: http://www.mentorfoundation.org/who_is_active.php?id=5 [accessed 2012-05-28] [WebCite Cache ID 67zxZgHMx]
16. Youth Voices Research Group. Youth Engagement + Health Promotion URL: <http://www.youthvoices.ca/> [accessed 2012-05-28] [WebCite Cache ID 67zxbKAYM]
17. Tercyak KP, Abraham AA, Graham AL, Wilson LD, Walker LR. Association of multiple behavioral risk factors with adolescents' willingness to engage in eHealth promotion. J Pediatr Psychol 2009 Jun;34(5):457-469. [doi: [10.1093/jpepsy/jsn085](#)] [Medline: [18723566](#)]
18. Skinner H, Biscope S, Poland B, Goldberg E. How adolescents use technology for health information: implications for health professionals from focus group studies. J Med Internet Res 2003 Dec 18;5(4):e32 [FREE Full text] [doi: [10.2196/jmir.5.4.e32](#)] [Medline: [14713660](#)]
19. Prokhorov AV, Kelder SH, Shegog R, Murray N, Peters R, Agurcia-Parker C, et al. Impact of A Smoking Prevention Interactive Experience (ASPIRE), an interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high-school students. Nicotine Tob Res 2008 Sep;10(9):1477-1485. [doi: [10.1080/14622200802323183](#)] [Medline: [19023839](#)]
20. Norman CD, Maley O, Li X, Skinner HA. Using the internet to assist smoking prevention and cessation in schools: a randomized, controlled trial. Health Psychol 2008 Nov;27(6):799-810. [doi: [10.1037/a0013105](#)] [Medline: [19025276](#)]

21. Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 1986 Sep;15(5):451-474. [Medline: [3534875](#)]
22. LeMaster PL, Connell CM, Mitchell CM, Manson SM. Tobacco use among American Indian adolescents: protective and risk factors. *J Adolesc Health* 2002 Jun;30(6):426-432. [Medline: [12039512](#)]
23. Murray DM, Pals SL, Blitstein JL. Design and analysis of group randomized trials in cancer prevention and control. In: Miller SM, Croyle RT, Rowland JH, editors. Miller SM, Bowen DJ, Croyle RT, Rowland JH, editors. *Handbook of Cancer Control and Behavioral Science: A Resource for Researchers, Practitioners, and Policymakers*. Washington, DC: American Psychological Association; 2008.
24. Sequist TD, Cullen T, Ayanian JZ. Information technology as a tool to improve the quality of American Indian health care. *Am J Public Health* 2005 Dec;95(12):2173-2179. [doi: [10.2105/AJPH.2004.052985](#)] [Medline: [16257947](#)]

Abbreviations

ASPIRE: A Smoking Prevention Interactive Experience

Edited by G Eysenbach; submitted 11.11.10; peer-reviewed by D Buller; comments to author 22.03.11; revised version received 23.06.11; accepted 09.03.12; published 01.06.12.

Please cite as:

Bowen DJ, Henderson PN, Harvill J, Buchwald D

Short-term Effects of a Smoking Prevention Website in American Indian Youth

J Med Internet Res 2012;14(3):e81

URL: <http://www.jmir.org/2012/3/e81/>

doi: [10.2196/jmir.1682](#)

PMID: [22659390](#)

©Deborah J. Bowen, Patricia Nez Henderson, Jessica Harvill, Dedra Buchwald. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Design of an mHealth App for the Self-management of Adolescent Type 1 Diabetes: A Pilot Study

Joseph A Cafazzo^{1,2,3}, PEng, MHSc, PhD; Mark Casselman¹, MSc; Nathaniel Hamming¹, PEng, MHSc; Debra K Katzman^{4,5}, MD, FRCPC; Mark R Palmert^{5,6}, MD, PhD

¹Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

²Institute of Health Policy, Management and Evaluation, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

³Institute of Biomaterials and Biomedical Engineering, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

⁴Division of Adolescent Medicine, The Hospital for Sick Children, Toronto, ON, Canada

⁵Department of Pediatrics, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

⁶Division of Endocrinology, The Hospital for Sick Children, Toronto, ON, Canada

Corresponding Author:

Joseph A Cafazzo, PEng, MHSc, PhD

Centre for Global eHealth Innovation

Techna Institute

University Health Network

Toronto General Hospital, R Fraser Elliott Building, 4th Floor

190 Elizabeth Street

Toronto, ON, M5G 2C4

Canada

Phone: 1 416 340 3634

Fax: 1 416 340 4955

Email: joe.cafazzo@uhn.ca

Abstract

Background: The use of mHealth apps has shown improved health outcomes in adult populations with type 2 diabetes mellitus. However, this has not been shown in the adolescent type 1 population, despite their predisposition to the use of technology. We hypothesized that a more tailored approach and a strong adherence mechanism is needed for this group.

Objective: To design, develop, and pilot an mHealth intervention for the management of type 1 diabetes in adolescents.

Methods: We interviewed adolescents with type 1 diabetes and their family caregivers. Design principles were derived from a thematic analysis of the interviews. User-centered design was then used to develop the mobile app *bant*. In the 12-week evaluation phase, a pilot group of 20 adolescents aged 12–16 years, with a glycated hemoglobin (HbA_{1c}) of between 8% and 10% was sampled. Each participant was supplied with the *bant* app running on an iPhone or iPod Touch and a LifeScan glucometer with a Bluetooth adapter for automated transfers to the app. The outcome measure was the average daily frequency of blood glucose measurement during the pilot compared with the preceding 12 weeks.

Results: Thematic analysis findings were the role of *data collecting rather than decision making*; the need for *fast, discrete transactions*; *overcoming decision inertia*; and the need for *ad hoc information sharing*. Design aspects of the resultant app emerged through the user-centered design process, including simple, automated transfer of glucometer readings; the use of a social community; and the concept of gamification, whereby routine behaviors and actions are rewarded in the form of iTunes music and apps. Blood glucose trend analysis was provided with immediate prompting of the participant to suggest both the cause and remedy of the adverse trend. The pilot evaluation showed that the daily average frequency of blood glucose measurement increased 50% (from 2.4 to 3.6 per day, $P = .006$, $n = 12$). A total of 161 rewards (average of 8 rewards each) were distributed to participants. Satisfaction was high, with 88% (14/16 participants) stating that they would continue to use the system. Demonstrating improvements in HbA_{1c} will require a properly powered study of sufficient duration.

Conclusions: This mHealth diabetes app with the use of gamification incentives showed an improvement in the frequency of blood glucose monitoring in adolescents with type 1 diabetes. Extending this to improved health outcomes will require the incentives to be tied not only to frequency of blood glucose monitoring but also to patient actions and decision making based on those readings such that glycemic control can be improved.

KEYWORDS

Type 1 diabetes mellitus; adolescent; cellular phone; self-care; chronic disease

Introduction

Chronic disease presents a growing challenge to the health and social care systems in Canada. More than 80% of primary care visits and two-thirds of medical admissions into hospital emergency departments are related to chronic diseases. Effective chronic disease management can result in improved health outcomes and increased quality of life [1].

Wagner [2] and others have proposed a practical framework for chronic disease management. Principally, it proposes that an “informed, activated patient” is needed to provide productive interactions with a “prepared, proactive practice team” [2]. As well, the Health Belief Model introduced the notion of creating a “cue to action” for the patient that elicits health behavior change [3]. Creating such conditions has been used in the design of many chronic disease management interventions across diverse disease populations. One group remains particularly elusive in creating conditions for positive health behavior change, that being adolescents with type 1 diabetes.

Type 1 diabetes mellitus is a chronic condition that is diagnosed in childhood and requires a lifetime of self-management at home and in the community between regular consultations with the health care team. Intensive self-management characterized by frequent self-monitoring of blood glucose is critical in type 1 diabetes to achieve good metabolic control. Improved blood glucose control has been shown to reduce mortality and the incidence of severe and costly complications such as renal and cardiovascular disease. Intensive self-management, including measurement of blood glucose at least 3 times per day, makes it theoretically possible to maintain near-normal blood glucose levels in patients with type 1 diabetes [4].

Despite its importance and a theoretical ability to optimize blood glucose control, worldwide data have repeatedly demonstrated that adolescents do not meet therapeutic targets [5]. Moreover, data from an international study group on childhood diabetes have not demonstrated a correlation between insulin regimen and glycemic control [6], suggesting that factors such as self-care behaviors and educational models likely have substantial impact on outcomes and that increased attention to these factors may lead to improved blood glucose control.

It has been well established that many struggle with interpreting and responding appropriately to the complex data sets that are part of effectively managing type 1 diabetes in real time [7]. Intensive self-management is a challenge for everyone with diabetes, but the mismatch between the demands of intensive self-management tasks and the developmental stage of adolescence often makes self-management even more difficult for this age group and may result in suboptimal glycemic control.

However, there is some recognition and willingness on the part of parents to use technology to assist in the management of their

child's diabetes, particularly when parents have felt that they have some unmet needs with respect to their care [8]. As well, given adolescents' propensity for new technology, such interventions may provide important opportunities to engage them and to help them improve self-management skills and behaviors [9]. A recent global survey indicated that adolescents around the world are adopting mobile technology faster and in a more immersive way than any previous generation. Recent Pew Internet & American Life Project reports indicate that the mobile phone has become the primary communication tool for the majority of adolescents in the United States; 75% of 12- to 17-year-olds now own mobile phones (up from 45% in 2004) [10]. Text messaging via mobile phone has become the most frequent form of adolescent interaction with friends (overtaking phone calls and face-to-face communication) [10].

The use of the telephone alone, however, has not had much clinical impact. A recent study showed how traditional cognitive behavior interventions such as educational sessions through phone calls had little value in changing negative health behaviors in children with type 1 diabetes [11]. An Australian study of 123 children with an average age of 11.9 years showed that a telephone call intervention by diabetes care providers biweekly did not result in improvements in glycated hemoglobin (HbA1c), diabetes knowledge, psychological parameters, or compliance [12]. A recent systematic review indicated that, although communication technologies may increase the frequency of contact between patient and health care professional, it remains unclear whether this results in improved outcomes [13].

More advanced telehealth methods have not fared better. In a study by Lehmkuhl and colleagues [14], the outcome improvements were unclear and data showed decreased engagement with parents. Youths in treatment reported increased unsupportive and decreased caring parental behavior, although the intervention did improve access to knowledgeable providers and resulted in a clinically significant improvement in glycemic control. There was, however, no significant difference when compared with the control group [14]. Further, a systematic review of type 1 diabetes mobile phone-based interventions indicated that the approach holds great promise, but few studies have shown definitive proof of improved health outcomes in this population [15].

In contrast, in the adult population with type 2 diabetes, a recent randomized controlled trial showed a 1.9% drop in HbA1c in the intervention group using a mobile phone-based remote monitoring and coaching system [16]. The technology used in this and other studies reflects the growing use information and communication technologies (eg, Internet, telephone, mobile phone, and Bluetooth) to track and transmit blood glucose results among adults with diabetes. A systematic review of 17 diabetes telemonitoring studies in adult populations has examined data quality aspects of telemonitoring, effect on clinical outcomes, and impact on behavior of patients and clinicians [17]. Limited work has been done in the development and evaluation of these

technologies to enhance self-care in adolescents with type 1 diabetes.

Self-monitoring of blood glucose is critical for effective self-care of type 1 diabetes, but adolescents with diabetes may also require decision-support aids to effectively contextualize a blood glucose result and take appropriate action to optimize glycemic control. The use of remote monitoring technology is likely not enough to elicit positive health behavior in this population. As well, the theoretical foundation of the behavioral intervention that is being delivered should be well established and proven. Heron and Smyth [18] recommended, based on their findings of a systematic review of ecological interventions for health behavior, that mobile technology-based ecological momentary interventions can be effectively implemented for a variety of health behaviors and psychological and physical symptoms. They also recommended the use of ecological momentary interventions that are dynamically and individually tailored and ecologically sensitive [18].

Conclusive methods of engaging the adolescent population in self-care remain experimental and elusive. The present study attempted to engage patients through the use of various approaches including reminders and cueing, social media communication, and the gamification of routine diabetes management tasks. As well, we designed this study to evaluate whether technology can be used to assist adolescents with self-care behaviors with a long-term objective of using technology to improve glycemic control among adolescents with type 1 diabetes.

This study engaged adolescents with type 1 diabetes, their families, and care providers in the design, development, and pilot evaluation of a home- and community-based diabetes telemanagement system. Given the ubiquity of mobile phones as a tool of daily living for many adolescents, we hypothesize that there may be a natural fit between the system and the target population. Hence, the intent of this study was to design, develop, and pilot an mHealth intervention for the management of type 1 diabetes in adolescent children.

Methods

We employed a user-centered design philosophy to gather requirements and iteratively design the system, leveraging a base remote patient monitoring system that has been tested in other populations. The previous-generation system used BlackBerry smartphones (Research In Motion, Waterloo, ON, Canada) and Bluetooth-enabled medical devices to capture client physiological information and generate and deliver alerts and reports to client, family, and provider [19–21]. Building on this experience, we describe the iterative, user-centered approach to design and develop a new iPhone-based (Apple Inc, Cupertino, CA, USA) system called *bant*.

User-Centered Design Phase

We employed user-centered design methods that are commonly used in contemporary design and, in particular, consumer-oriented products. This entails the iterative involvement of the end user in the design process by eliciting formal feedback on reference and prototype versions of the

intervention; heuristic evaluation by human factors specialists; and formative usability testing of the system.

This design phase involved qualitative, ethnographic interviews with patients and family caregivers (parents) to inform the design and development of the app. In addition, we held focus group sessions with a cross-section of clinical team representatives from the Division of Adolescent Medicine and the Diabetes Program in the Division of Endocrinology at the Hospital for Sick Children (SickKids) in Toronto. Patients aged 12–16 years receiving care at the Diabetes Clinic at SickKids were eligible if they had a diagnosis of type 1 diabetes for more than 1 year; could read, speak, and understand English; and were willing to participate in the study. Each adolescent participant was interviewed individually. Participant parents waited to be interviewed individually after their child. Finally, parent and child were interviewed together to complete the session.

We used an ethnographic research approach when conducting the interviews. Once a sufficient level of saturation was achieved through the focus groups, the recordings were transcribed verbatim. A general inductive method was used in the analysis of the transcripts. Transcripts were read repeatedly and text segments coded for potential themes. As the coding framework developed, transcripts were reanalyzed in light of new themes that may have emerged as a result. Once we completed this step, we derived major themes that were relevant to the research question. Coding was free not to assume any presuppositions. Themes that emerged were used to inform the design and development of the self-management system. A solitary reviewer analyzed the coding, which was validated through member checking of adolescent health and endocrinology specialists at SickKids Hospital.

Data saturation was achieved on completion of the sixth set of patient and parent semistructured interviews. Each interview was conducted by the research coordinator and guided by an interview guide that was based on the study objectives and a priori knowledge of diabetes management, behavior change theory, and health care software design.

Clinical Pilot

We chose a convenience sample of 20 adolescents to test the assumptions of the intervention. Patients aged 12–16 years receiving care at the Diabetes Clinic at SickKids were eligible if they had a diagnosis of type 1 diabetes for more than 1 year; had received care at the Diabetes Clinic at SickKids for more than 6 months; could read, speak, and understand English; and were willing to participate in the study. Inclusion also required an HbA1c of between 8% and 10% at the time of the previous clinic visit. Information letters describing the study were sent to patients who met the inclusion criteria. The letters identified a mechanism for patients who wanted to opt out so that they would not be approached for recruitment at the time of their next visit. The clinic staff identified the patients to the study project manager at the time of the visit for recruitment. If inclusion criteria were met, and patient and parent agreed to participate, informed consent was obtained. Of the 20 participants, 15 adolescents were given an iPhone 4 to use, and 5 were given an iPod Touch (fourth generation) to use. We felt

that this sample of 15 was sufficient to test the assumptions made in this initial iteration of the intervention. The iPod Touch group was added to test whether having a Wi-Fi-only device would affect the user's experience of the intervention or would alter the patterns of use.

The protocol was approved by the Hospital for Sick Children's Research Ethics Board (1000017742). As well, we obtained an Investigational Test Authorization (ITA) through Health Canada, due to the advanced technical integration of the blood glucometer with the iPhone app. This was to ensure from a regulatory perspective that the medical device used in the trial was safe. Patients were trained on the system after they consented and were asked to use the system until the time of their next clinic visit, in approximately 12 weeks.

The primary outcome measure was frequency of daily blood glucose readings. The baseline measure was established based on a data extraction of the patient's personal blood glucose meter at the time of recruitment. We performed a paired *t* test comparing the average number of blood glucose readings 3 months prior to recruitment (baseline) and during the approximately 12 weeks of use of the system during the pilot. HbA_{1c} was also measured at baseline and at the time of the following clinic visit.

We selected well-established measures of diabetes-related self-efficacy, self-care behavior, quality of life, and adherence to self-care psychosocial variables, all of which have well-documented reliability and validity in this population as provided in the references following. Self-care behavior and treatment adherence were assessed at baseline and postintervention using the validated 14-item Self-Care Inventory [22]. Parent-adolescent interaction around diabetes care tasks and decision making were assessed at baseline and postintervention using the Diabetes Family Responsibility Questionnaire [23]. Diabetes-specific quality of life was assessed at baseline and postintervention using the Diabetes Quality of Life for Youth instrument [24].

The study coordinator provided surveys in paper form to participants. Participants completed the survey forms in a secluded area of the diabetes clinic waiting room while they

waited for their appointment with their care team. The study coordinator was available in case participants had any questions about the surveys.

Results

User-Centered Design Phase

We interviewed 6 adolescents (with 1 parent for each) for this design phase to achieve saturation.

The following major themes emerged and were used to inform the design and development of the self-management system.

Theme: Fast, Discrete Transactions

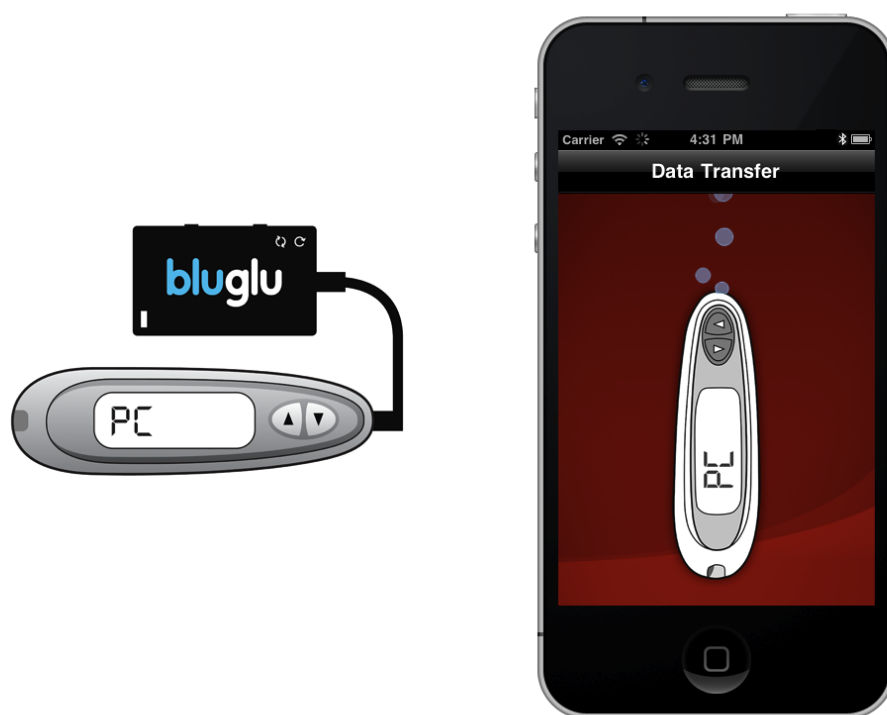
The adolescent participants noted the need for fast transactions taking seconds, not minutes. Several participants commented that social embarrassment was a key factor leading to their avoidance of testing in public (eg, skipping the lunch test in a busy school cafeteria).

Design Principle

The design principles for fast, discrete transactions were to (1) ensure that interactions with the system would be fast, (2) use wizards (algorithms used to provide prompts based on available data) to guide user interaction where possible, (3) design the intervention to fit the adolescent lifestyle, and (4) make the interactions related to diabetes discrete.

To achieve this goal, we elected to have automated data transfers from their glucometer, rather than having participants enter data manually. We designed an adapter (called *bluglu*), which allows a OneTouch UltraMini glucometer (LifeScan, Inc., Milpitas, CA, USA) to communicate via Bluetooth, allowing the transfer of blood glucose reading wirelessly, to the iPhone device running *bant* (Figure 1). Data can be transferred from the meter to *bant* at any time. If readings are taken at a time of day (eg, during school lunch) when the iPhone device is not available, data can be transferred to *bant* later in the day at the user's convenience. However, as soon as the data are transferred, the analysis tools assess the data so that the adolescent gets feedback in real time.

Figure 1. The bluglu adapter for wireless transfer of blood glucose readings via Bluetooth. This avoids the need for manual data entry by the user.



Theme: The Role of Data Collecting Rather Than Decision Making

Several participants noted that they did not typically use tracking and analysis tools to review their test results because all of the information was “in the meter.” There seemed to be tension between the roles of tester/collector and analyst/decision maker.

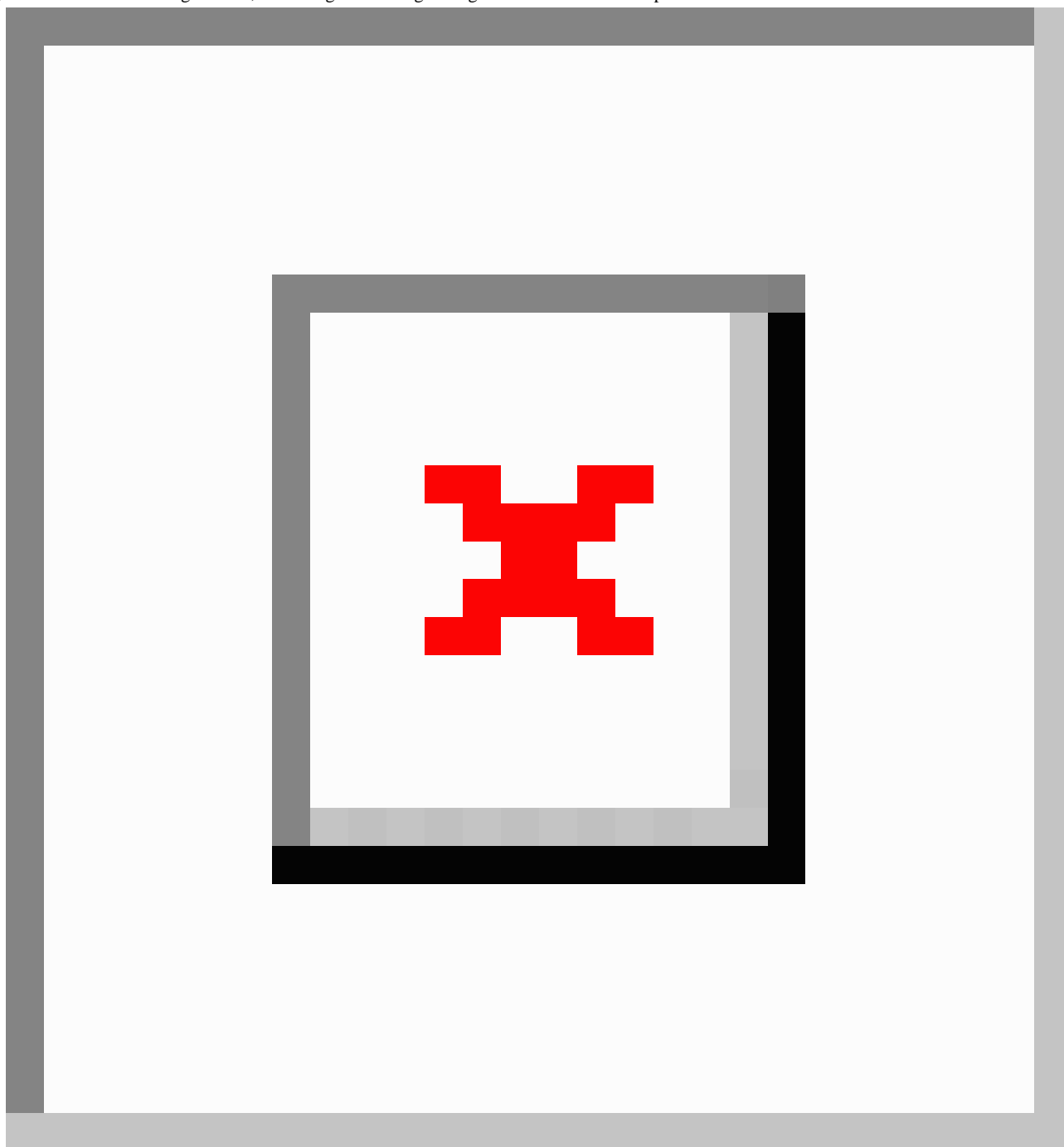
Design Principle

Simple data display and decision-support prompts and alerts that integrate into the daily workflow of blood glucose testing may help adolescents to take on more analysis and

decision-making tasks in a timely manner, leading to more proactive management.

To address this design principle, we endeavored to give the data greater context and value for decision making. Hence, we designed a novel visual display that has been validated by human factors experts, who specialize in optimizing human–computer interaction, clinicians, app design specialists, and adolescents with diabetes. This provides the user with a summary of daily glycemic control at a glance, associating each data point with context, related to meals and activity, and highlights when blood glucose values are out of range (Figure 2).

Figure 2. The *bant* Readings screen, indicating the reading through colored circles that represent both value and context.



Theme: Overcoming Decision Inertia

In some cases, adolescents were making few changes to their daily diabetes regimen, even when the blood glucose data had been collected and the profile suggested a change was needed.

Design Principle

The design principles to overcome decision inertia were to (1) help adolescents identify blood glucose trends, (2) promote cognitive processing related to identifying and correcting the trend, and (3) integrate rewards and incentives into the system to sustain engagement with the tool.

When blood glucose levels are out of range for 3 days in a row in a particular context (eg, before breakfast), *bant* detects the

3-day trend and prompts the user to make a decision about the cause of the trend and how to rectify it. In addition, *bant* provides data analysis and trending screens that display the percentage of blood glucose levels that are in range at specific times (eg, before meals, before bedtime, or overnight) (Figure 3) as well as decision support via the trend wizard feature, which assists the adolescent in identifying the cause of the trend and the adjustment in regimen that might help improve blood glucose control (Figure 4).

To further incentivize adolescents to use the app, we designed *bant* with a rewards algorithm that allocated gamelike experience points for adhering to best-practice guidelines for blood glucose testing (goal of three or more tests per day). The algorithm provided increasing point allocations for each test

performed (maximum points were awarded for five blood glucose tests performed across five different contexts). Points were awarded based on the instances of each reading and consecutive readings; and bonus points were awarded for a full day of readings (Figure 5). “Leveling-up” was achieved when users had earned 200 experience points, whereby they could be redeemed for Apple iTunes and App Store purchases, usually in Can \$1 increments, so the rewards were frequent, up to one reward every 2–3 days for fully adherent users.

As well, we added the ability for users to communicate with their peers in a secure community area of the app to share experiences and gain or provide support (Figure 5). This was achieved through a private microblogging platform similar to the social network Twitter. It used the open-source alternative StatusNet (available at <http://status.net/>) running on a secure server at our center at Toronto General Hospital. The intent here was for the social aspect of the app to increase regular use of the app and the positive health behaviors it was intended to elicit.

Figure 3. Trends allows at-a-glance review of readings over 7-day, 14-day, 21-day, and 90-day periods. The percentages of readings under, within, and above target range are noted for quick review.

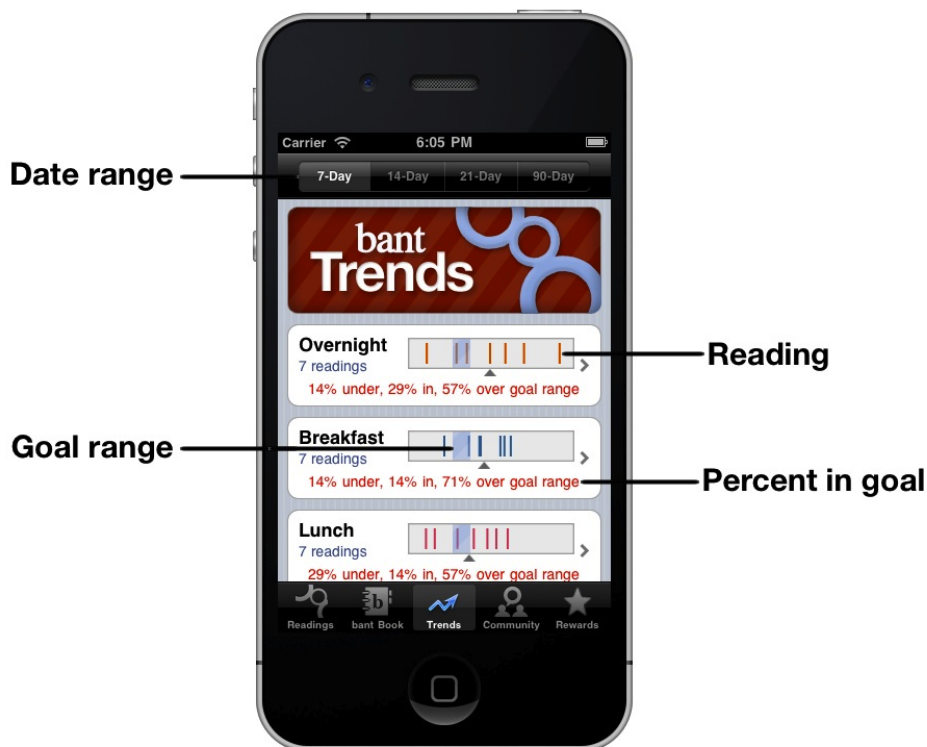


Figure 4. Trend Alert automatically identifies 3-day trends and prompts users to identify what they believe is the cause (left) and their intended action (right).

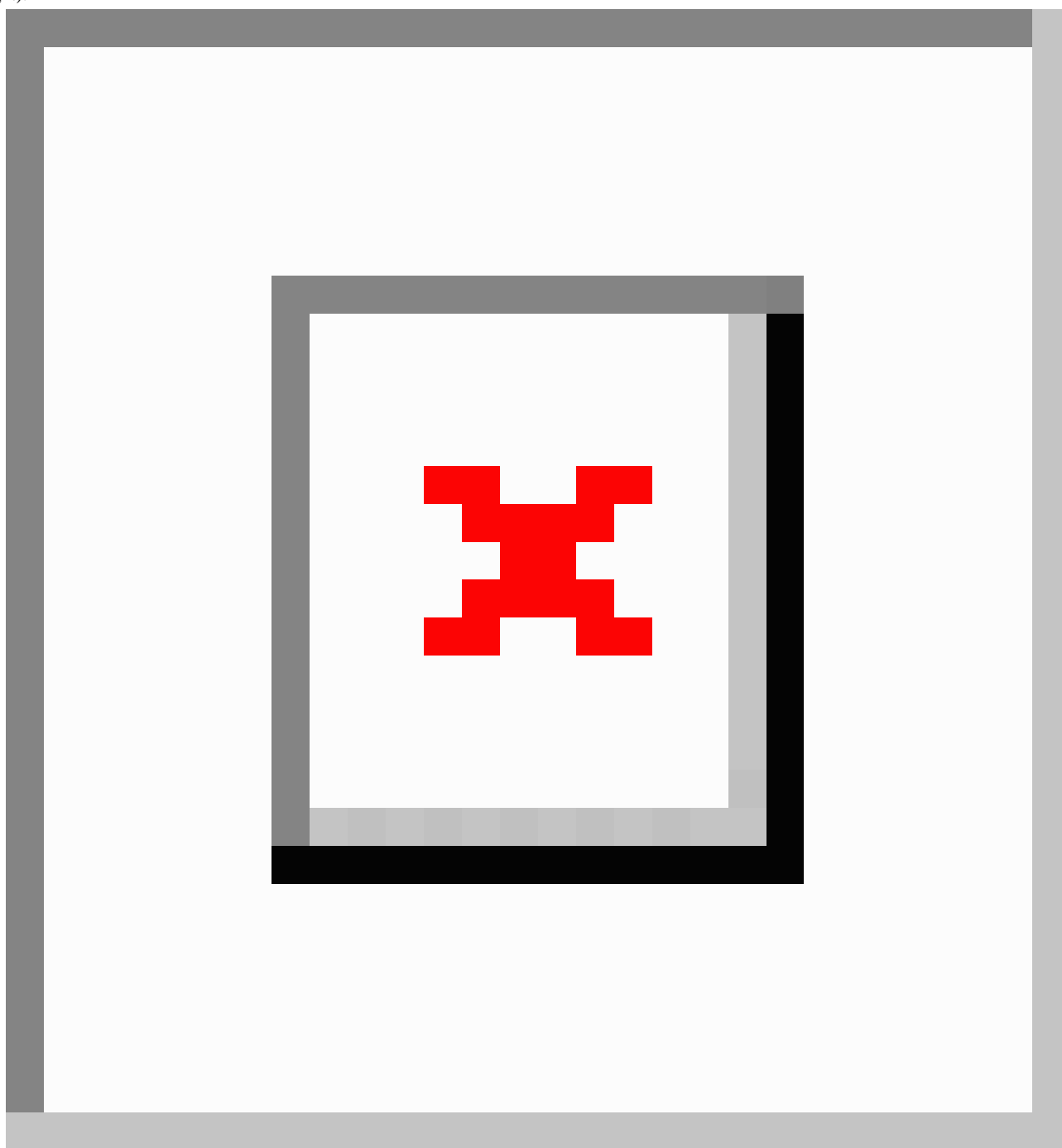
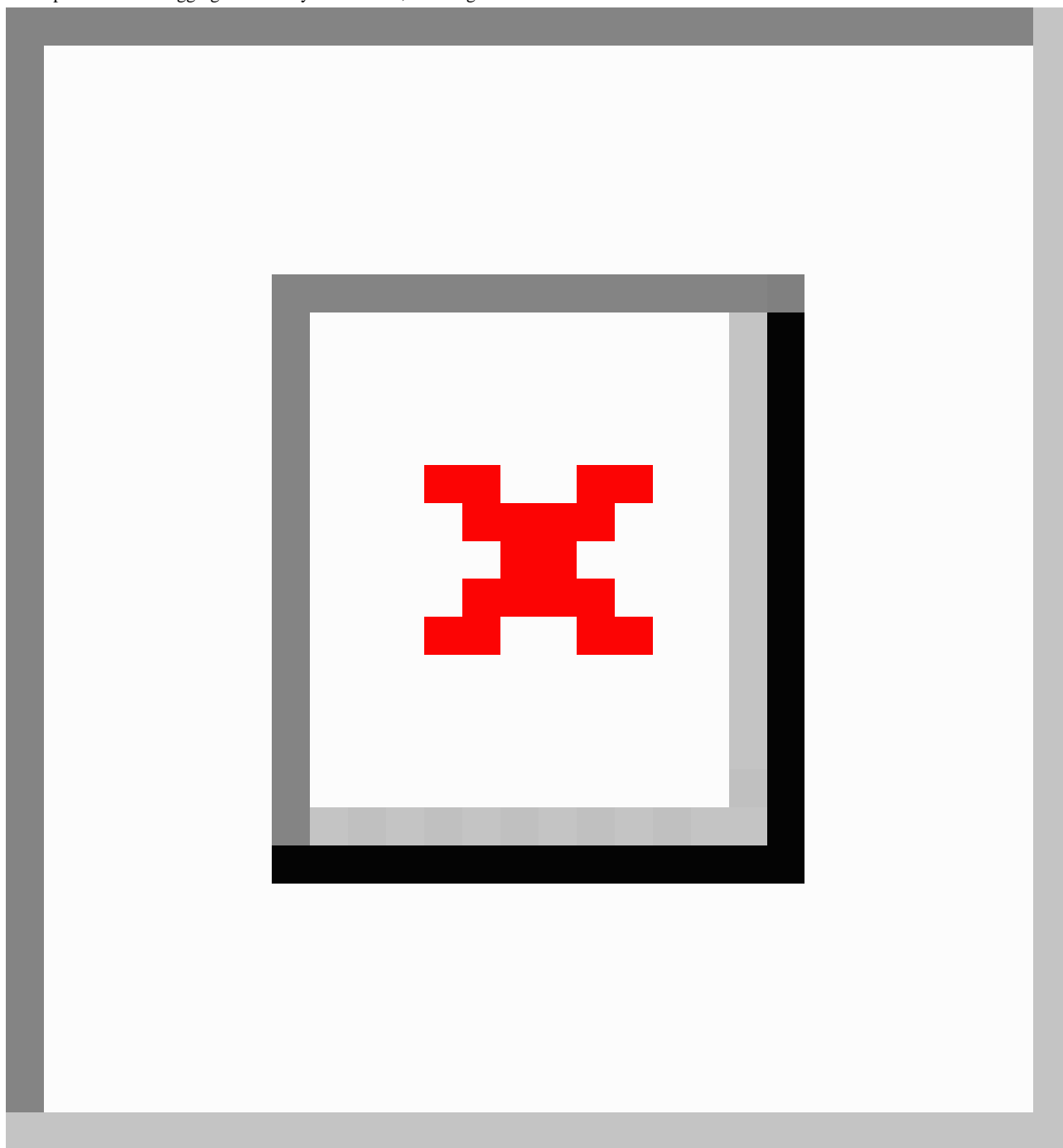


Figure 5. The *bant* Rewards screen (left) indicating experience points (XP) awarded and redeemed for applications and music. The right-hand screen shows the private microblogging community within *bant*, allowing users to communicate with one another.



Theme: Ad Hoc Information Sharing

Several adolescent participants and parents described an ad hoc approach to day-to-day sharing of diabetes-related information, including test results. Teens generally expressed a willingness to share results with parents and valued their input when it was requested.

Design Principle

Ad hoc information sharing would provide opportunities for adolescents to safely share test results and diabetes-related information with parents (as well as peers and clinic staff) via secure online tools and communities for sharing results.

To address this design principle, we elected to integrate the mobile app with a secure online personal health record called TELUS health space, a localized version of Microsoft HealthVault, addressing the needs of Canadian health care practice, including spoken and written language and units. User data could be stored securely and optionally shared with parents and providers. This personal health record is a precertified consumer platform approved by the federal agency Canada Health Infoway. Both patients and providers were provided with health space accounts so as to be able to view shared data from the adolescent.

Clinical Pilot

The average age of the participants was 14.9 years (SD 1.3 years, $N = 20$), equally represented by gender (10 each). Of the participants, 9 used an insulin pump and 11 used multiple daily injections.

Principal Result: Daily Average Frequency of Blood Glucose Measurement

We recruited 20 patients as planned. Of these, 2 did not complete the full 12-week pilot study and 6 did not have sufficient baseline data from the meters for us to perform the analysis for the primary outcome measure. Hence, we analyzed change in daily frequency of measurements for the remaining 12 patients. For this subgroup analysis ($n = 12$), the average age of the participants was 15.1 (SD 1.3) years, with 8 girls and 4 boys. Of these participants, 8 used an insulin pump and 4 used multiple daily injections.

Daily average frequency of blood glucose measurement increased 50% (from 2.4 readings per day to 3.6 readings per day, $P = .006$).

Principal Result: Rewards

The level of patient engagement in the rewards program indicated the effectiveness of the rewards feature. A total of 161 rewards (an average of 8 rewards each, $N = 20$) were distributed to patients based on their frequency of measurement, with 50% (10/20) of patients collecting more than 10 awards, 25% (5/20) collecting fewer than 10 rewards, and 25% (5/20) collecting no rewards. There were 2 patients who accumulated a significant number of points and were highly adherent but never redeemed the points for rewards, indicating that the form of reward is not the sole motivator for some patients.

Secondary result: HbA1c

HbA1c did not change significantly over the pilot period (average 8.8%, SD 0.74 vs 9.2%, SD 1.03, $P = .11$).

Results of Self-Reported Measures

Self-care Inventory

We calculated results of the 14-item survey of how well the adolescents followed their prescribed self-care regimen from complete responses from 14 of the 20 participants at baseline and the conclusion of the pilot. The instrument uses a 5-point Likert scale ranging from 1 (never do it) to 5 (always do this as recommended without fail). We found no significant changes in the dimensions of adherence (pretrial average score 3.5, SD 0.93, vs posttrial average 3.6, SD 0.93), blood glucose regulation (average 3.6, SD 0.91 vs 3.6, SD 0.99), insulin and food regulation (average 3.5, SD 0.91 vs 3.6, SD 0.81), and emergency preparedness (average 3.4, SD 1.47 vs 3.4, SD 1.54). There was evidence of improvement in exercise (average 2.9, SD 1.09 vs 3.5, SD 1.15), but it was still not significant ($P = .069$). This finding could be attributed to the seasonal effect of the pilot beginning in the winter and ending in the early spring. The intervention had no specific design aspect for exercise.

Diabetes Family Responsibility Questionnaire

Parent–adolescent patient interaction around diabetes care tasks and decision making was assessed at baseline and postintervention. Data for 14 of 20 patients were available. We found slight improvements in both the parent and adolescent scores. None were significant. The average score from the caregivers' perspective improved from 1.9 (SD 0.14) to 2.0 (SD 0.16). The average score from the adolescents' perspective improved from 2.1 (SD 0.14) to 2.2 (SD 0.15). A score of 1 indicates the caregiver initiates the responsibility and a score of 3 indicates the child initiates the responsibility. A score of 2 indicates a shared responsibility.

Diabetes Quality of Life Instrument for Youth

The 22-item survey results were calculated from complete responses from 14 of 20 patients at baseline and the conclusion of the pilot. We found no significant changes in the dimensions of impact of symptoms (pretrial average 3.4, SD 1.06 vs posttrial average 3.9, SD 1.76), impact of treatment (average 3.7, SD 1.77 vs 3.6, SD 1.33), impact of activities (average 2.9, SD 2.4 vs 3.5, SD 0.24), parent issues (average 8.1, SD 2.71 vs 8.5, SD 2.62), worries about diabetes (average 7.5, SD 4.32 vs 9.0, SD 6.05), and health perception (average 2.2, SD 0.66 vs 2.3, SD 0.68). Generally, most of the dimensions showed either no change or a trend toward worsening in the quality of life.

Social Networking Findings

A total of 288 posts were made to the community over the pilot period ($N = 20$). The average was 14 posts per patient over the period. The results were highly skewed with a median of 4.5 and interquartile range of 12.5. As is the case with social networks, a minority of patients posted most frequently: 5 of 20 patients posted 53.1% (153/288) of all posts, and 3 posted only once. Most of the posts were categorized as being either social in nature, accounting for 43.1% (124/288) of the total, or related to the pilot study itself (127/288, 44.1% of the total). The remaining 12.8% (37/288) of posts were categorized as being medically related, including posed questions (12/288, 4%), answers (12/288, 4%), or comments (13/288, 5%).

Perceptions and Satisfaction

Satisfaction was high, with 88% (14/16 of those who participated in poststudy exit interviews) stating that they would continue to use the system. The remaining 2 patients requested that the app be integrated with their insulin pump (or use a meter that communicates directly with the pump) before continuing to use the system.

Discussion

These findings revealed specific requirements that were expressed as four themes: the need for *fast, discrete transactions*, the role of *data collection rather than decision making*, overcoming *decision inertia*, and *ad hoc information sharing*. The pilot trial of the resultant intervention showed an improvement in the daily average frequency of blood glucose measurements by 50%. HbA1c did not change significantly. Other self-reported indicators showed no significant results. Satisfaction with the intervention was high, with 88% (14/16 participants) stating that they would continue to use the system.

In the past, several information and communication technologies (eg, Internet, telephone, mobile phone, and Bluetooth-connected blood glucose meters) have been used to track and transmit blood glucose results. However, many of these earlier interventions were developed as electronic logbook applications to collect, store, and display blood glucose readings. Few featured interactive prompts that, as an example, would cue adolescents with diabetes to take action and adjust their treatment protocol. A systematic review of telemedicine interventions to support self-monitoring of blood glucose noted that few studies have examined the relationship between capture of blood glucose data, analysis, advice, and subsequent behavior change. The authors suggest that self-monitoring of blood glucose “with or without telemedicine is only likely to be helpful when test results are linked to educational or behavioural advice and changes in clinical management” [25]. Limited work has been completed on home- and community-based telemonitoring or telemanagement to enhance self-care in adolescents with type 1 diabetes [26]. Furthermore, the authors of this review identified that none of the electronic interventions studied have been particularly effective and that a new approach is needed.

Hence, our approach to the design was not only to capture the end-user requirements of adolescents and parents in managing type 1 diabetes, but also to test new concepts of behavior change. In particular, the use of social networking and rewards were attempts to create adherence to the use of the app, as well as elicit positive health behavior.

The primary outcome of the pilot study was achieved with a significant increase in the daily frequency of blood glucose readings. More frequent self-monitoring of blood glucose (≥ 3 times daily) is associated with better glycemic control among patients with type 1 diabetes [27]. We hypothesize that the behavioral mechanism that produced this significant result was the combination of the simple automated reminders that were set for the patients at the start of the trial and the issuing of rewards tied directly to taking readings. The use of the rewards system and its apparent effect on eliciting positive health behavior is an example of gamification, which is commonly used in other business sectors, such as commercial patronage loyalty programs. Its application to health care is relatively new. Although games have been used in previous health care behavioral interventions, few examples have demonstrated effectiveness. Rather than creating a diabetes-themed game, we believed that the use of gamelike features of routine self-management tasks would have a greater likelihood of success and could be sustained over longer periods.

For the secondary outcome, however, HbA1c did not change significantly over the pilot period. This result was not entirely unexpected given the small sample size, the short duration of the pilot, and the chronic management problems of the participants involved. Nonetheless, the purpose of the intervention was to elicit positive health behaviors by increasing the frequency of daily blood glucose readings, and we incentivized this behavior, not making changes in the treatment regimen or decreasing the number of out-of-range blood glucose values. In our study, although behavior change related to blood glucose monitoring was achieved, it did not result in improvements in HbA1c; however, previous studies have shown

that blood glucose measurement frequency was significantly associated with better metabolic control, with a drop of HbA1c of 0.2% for one additional reading per day ($P < .001$) [28]. We hypothesize that in a properly powered study of sufficient duration that incentivizes not only blood glucose measurements but also metrics of improved glycemic control, HbA1c would improve.

The use of the Trend Alerts aspect of the app is an example of what Heron and Smyth referred to as an ecological momentary intervention [18]. This feature did not show any direct value to health outcomes, however, but we hypothesize that tying reward incentives to actions related to the Trend Alerts may improve health outcomes.

There were no definitive findings from the self-reporting instruments. However, there is some indication that interventions such as *bant* can increase anxiety among patients. This has been observed in other studies by Logan et al in the adult hypertension population [20]. This creates a paradox in that we are attempting to elicit behaviors in patients that will make them more mindful of their condition, yet may have the added negative effect of creating anxiety about their condition. Clearly, the design of the interventions must be cognizant of the potential negative effects, and extra care must be taken in the form of the messaging that is provided to the patient.

We were pleased to learn that schools were very accommodating of the participants when informed of the purpose of the smartphones. In fact, when one was stolen from the locker room, a school principal made an announcement for its return. It was recovered shortly thereafter. No personal information was breached, as a remote wipe command was issued through a device management server when the loss was reported. The device was recovered in its default state.

Limitations

The pilot was limited to a single-site, small convenience sample with no control group, which limited our ability to generalize the findings and determine the efficacy of the intervention fully. Importantly, the one positive outcome was directly related to the rewards incentive and elicited the expected results. Future iterations of the intervention will attempt to apply this approach to other aspects of patient self-care, leading to a randomized controlled trial.

Comparison with Prior Work

A recent study of the physically integrated phone and glucose meter GlucoPhone (HealthPia, Paducah, KY, USA) in 40 adolescents with type 1 diabetes showed positive feelings toward the intervention despite having significant technical issues during the 6-month trial [29]. It did not show any improvement in the participants' quality of life, level of conflict with parents, reported self-management, or glycemic control. This outcome occurred despite significant improvements in health outcomes in the adult type 2 population with the same technology platform previously [30]. Carroll and colleagues speculated that the outcomes with the adolescents were a result of a lack of explicitly attempting to change health behavior. They suggest that adding behavioral contracts could improve the chances of better outcomes [29].

A follow-up study by the same authors showed the additional use of a behavioral contract to be effective in changing the behavior of adolescents with type 1 diabetes [31]. The contract was in the form of a written pledge that the adolescent signed in agreement with the terms set out by the caregiver. The investigators incentivized the adolescent to practice self-care in return for less nagging from the caregiver. The 3-month study was with 10 adolescent-caregiver pairs. Outcomes were improvement in the diabetes self-management profile and significant improvement in HbA1c. However, it is unclear how the behavioral contract is tied to the technological intervention and whether the GlucoPhone was even needed to achieve such a result, other than monitoring adherence to the agreed terms.

Hence, we believe that for the technological intervention to have intrinsic value, the behavioral aspects must be part of the app itself. As nascent as this area of research is, we have found the use of rewards to be effective.

Conclusions

The use of mHealth behavioral interventions is a new area of research, where there is little data to show whether and how

such technology can be used effectively. In our pilot, we demonstrated the feasibility of deploying a diabetes app for use with the adolescent type 1 diabetes population. We also showed that the use of gamification incentives was associated with an improvement in the frequency of blood glucose monitoring in this group. Extending the use of *bant* to improve health outcomes may require that the incentives be tied not only to frequency of blood glucose monitoring but also to patient actions and decision making based on those readings, such that glycemic control can be improved. The long-term sustainability of using iTunes redemption codes as a reward is an issue that requires attention. We would encourage experimentation related to the use of nontangible, virtual rewards as a substitute.

We also have few findings on the effect of the community aspect of the intervention, and its direct value is difficult to ascertain. Only through wider deployment can the use of these social networking aspects be properly evaluated.

Although we cannot fully generalize these results without a control group trial, the findings indicate that the use of these design principles show promise in eliciting positive health behaviors in adolescents with type 1 diabetes.

Acknowledgments

The authors wish to thank the patients and clinicians of the Diabetes Clinic of the Hospital for Sick Children for participating in the pilot.

This study was partially funded by the Applied Client-focused Team Research Grants Program of Saint Elizabeth Health Care.

We also acknowledge the contributions of the team that developed the system *bant* from the Centre for Global eHealth Innovation: Kevin Tallevi, Mathieu Trudel, Peter Picton, Archana Gopal, and Mark Bianco. As well, we thank the members of the University Health Network Healthcare Human Factors team for testing and quality control: Tara McCurdie, Deborah Chan, Wayne Ho, and Anjum Chagpar.

Conflicts of Interest

None declared. TELUS Health contributed the iPhone 4s, voice and data service, and TELUS health space accounts for the pilot. All other equipment was supplied through grant funding or internally sourced.

References

1. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Med Care* 1999 Jan;37(1):5-14. [Medline: [10413387](#)]
2. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Eff Clin Pract* 1998;1(1):2-4 [FREE Full text] [Medline: [10345255](#)]
3. Becker MH. The Health Belief Model and Personal Health Behavior. Thorofare, NJ: CB Slack; 1974.
4. Anonymous. Effect of intensive diabetes treatment on the development and progression of long-term complications in adolescents with insulin-dependent diabetes mellitus: Diabetes Control and Complications Trial. Diabetes Control and Complications Trial Research Group. *J Pediatr* 1994 Aug;125(2):177-188. [Medline: [8040759](#)]
5. Holl RW, Swift PG, Mortensen HB, Lynggaard H, Hougaard P, Aanstoot HJ, et al. Insulin injection regimens and metabolic control in an international survey of adolescents with type 1 diabetes over 3 years: results from the Hvidoere study group. *Eur J Pediatr* 2003 Jan;162(1):22-29. [doi: [10.1007/s00431-002-1037-2](#)] [Medline: [12486503](#)]
6. de Beaufort CE, Swift PG, Skinner CT, Aanstoot HJ, Aman J, Cameron F, Hvidoere Study Group on Childhood Diabetes 2005. Continuing stability of center differences in pediatric diabetes care: do advances in diabetes treatment improve outcome? The Hvidoere Study Group on Childhood Diabetes. *Diabetes Care* 2007 Sep;30(9):2245-2250. [doi: [10.2337/dc07-0475](#)] [Medline: [17540955](#)]
7. Hood KK, Peterson CM, Rohan JM, Drotar D. Association between adherence and glycemic control in pediatric type 1 diabetes: a meta-analysis. *Pediatrics* 2009 Dec;124(6):e1171-e1179 [FREE Full text] [doi: [10.1542/peds.2009-0207](#)] [Medline: [19884476](#)]

8. Pena V, Watson AJ, Kvedar JC, Grant RW. Mobile phone technology for children with type 1 and type 2 diabetes: a parent survey. *J Diabetes Sci Technol* 2009 Nov;3(6):1481-1489 [FREE Full text] [Medline: [20144404](#)]
9. Skinner H, Biscope S, Poland B, Goldberg E. How adolescents use technology for health information: implications for health professionals from focus group studies. *J Med Internet Res* 2003 Dec 18;5(4):e32 [FREE Full text] [doi: [10.2196/jmir.5.4.e32](#)] [Medline: [14713660](#)]
10. Lenhart A, Ling R, Campbell S, Purcell K. Teens and mobile phones. Washington, DC: Pew Internet & American Life Project; 2010 20. URL: <http://www.pewinternet.org/~media/Files/Reports/2010/PIP-Teens-and-Mobile-2010-with-topline.pdf> [accessed 2012-04-23] [WebCite Cache ID 679DGWFdk]
11. Salamon KS, Hains AA, Fleischman KM, Davies WH, Kichler J. Improving adherence in social situations for adolescents with type 1 diabetes mellitus (T1DM): a pilot study. *Prim Care Diabetes* 2010 Apr;4(1):47-55. [doi: [10.1016/j.pcd.2009.10.003](#)] [Medline: [19926354](#)]
12. Nunn E, King B, Smart C, Anderson D. A randomized controlled trial of telephone calls to young patients with poorly controlled type 1 diabetes. *Pediatr Diabetes* 2006 Oct;7(5):254-259. [doi: [10.1111/j.1399-5448.2006.00200.x](#)] [Medline: [17054446](#)]
13. Sutcliffe P, Martin S, Sturt J, Powell J, Griffiths F, Adams A, et al. Systematic review of communication technologies to promote access and engagement of young people with diabetes into healthcare. *BMC Endocr Disord* 2011;11:1 [FREE Full text] [doi: [10.1186/1472-6823-11-1](#)] [Medline: [21210964](#)]
14. Lehmkuhl HD, Storch EA, Cammarata C, Meyer K, Rahman O, Silverstein J, et al. Telehealth behavior therapy for the management of type 1 diabetes in adolescents. *J Diabetes Sci Technol* 2010 Jan;4(1):199-208 [FREE Full text] [Medline: [20167185](#)]
15. Russell-Minda E, Jutai J, Speechley M, Bradley K, Chudyk A, Petrella R. Health technologies for monitoring and managing diabetes: a systematic review. *J Diabetes Sci Technol* 2009 Nov;3(6):1460-1471 [FREE Full text] [Medline: [20144402](#)]
16. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care* 2011 Sep;34(9):1934-1942. [doi: [10.2337/dc11-0366](#)] [Medline: [21788632](#)]
17. Jaana M, Paré G. Home telemonitoring of patients with diabetes: a systematic assessment of observed effects. *J Eval Clin Pract* 2007 Apr;13(2):242-253. [doi: [10.1111/j.1365-2753.2006.00686.x](#)] [Medline: [17378871](#)]
18. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol* 2010 Feb;15(Pt 1):1-39. [doi: [10.1348/135910709X466063](#)] [Medline: [19646331](#)]
19. Halifax NV, Cafazzo JA, Irvine MJ, Hamill M, Rizo CA, McIssac WJ, et al. Telemanagement of hypertension: a qualitative assessment of patient and physician preferences. *Can J Cardiol* 2007 May 15;23(7):591-594. [Medline: [17534469](#)]
20. Logan AG, McIsaac WJ, Tisler A, Irvine MJ, Saunders A, Dunai A, et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. *Am J Hypertens* 2007 Sep;20(9):942-948. [doi: [10.1016/j.amjhyper.2007.03.020](#)] [Medline: [17765133](#)]
21. Cafazzo JA, Seto E. The hospital at home: advances in remote patient monitoring. *Biomed Instrum Technol* 2010;Suppl Home Healthcare:47-52. [Medline: [22049607](#)]
22. Lewin AB, LaGreca AM, Geffken GR, Williams LB, Duke DC, Storch EA, et al. Validity and reliability of an adolescent and parent rating scale of type 1 diabetes adherence behaviors: the Self-Care Inventory (SCI). *J Pediatr Psychol* 2009 Oct;34(9):999-1007. [doi: [10.1093/jpepsy/jsp032](#)] [Medline: [19423660](#)]
23. Anderson BJ, Auslander WF, Jung KC, Miller JP, Santiago JV. Assessing family sharing of diabetes responsibilities. *J Pediatr Psychol* 1990 Aug;15(4):477-492. [Medline: [2258796](#)]
24. Ingersoll GM, Marrero DG. A modified quality-of-life measure for youths: psychometric properties. *Diabetes Educ* 1991;17(2):114-118. [Medline: [1995281](#)]
25. Farmer A, Gibson OJ, Tarassenko L, Neil A. A systematic review of telemedicine interventions to support blood glucose self-monitoring in diabetes. *Diabet Med* 2005 Oct;22(10):1372-1378. [doi: [10.1111/j.1464-5491.2005.01627.x](#)] [Medline: [16176199](#)]
26. Shulman RM, O'Gorman CS, Palmert MR. The impact of telemedicine interventions involving routine transmission of blood glucose data with clinician feedback on metabolic control in youth with type 1 diabetes: a systematic review and meta-analysis. *Int J Pediatr Endocrinol* 2010;2010 [FREE Full text] [doi: [10.1155/2010/536957](#)] [Medline: [20886054](#)]
27. Karter AJ, Ackerson LM, Darbinian JA, D'Agostino RB, Ferrara A, Liu J, et al. Self-monitoring of blood glucose levels and glycemic control: the Northern California Kaiser Permanente Diabetes registry. *Am J Med* 2001 Jul;111(1):1-9. [Medline: [11448654](#)]
28. Ziegler R, Heidmann B, Hilgard D, Hofer S, Rosenbauer J, Holl R, DPV-Wiss-Initiative. Frequency of SMBG correlates with HbA1c and acute complications in children and adolescents with type 1 diabetes. *Pediatr Diabetes* 2011 Feb;12(1):11-17. [doi: [10.1111/j.1399-5448.2010.00650.x](#)] [Medline: [20337978](#)]
29. Carroll AE, DiMeglio LA, Stein S, Marrero DG. Using a cell phone-based glucose monitoring system for adolescent diabetes management. *Diabetes Educ* 2011;37(1):59-66. [doi: [10.1177/0145721710387163](#)] [Medline: [21106908](#)]

30. Cho JH, Lee HC, Lim DJ, Kwon HS, Yoon KH. Mobile communication using a mobile phone with a glucometer for glucose control in Type 2 patients with diabetes: as effective as an Internet-based glucose monitoring system. *J Telemed Telecare* 2009;15(2):77-82. [doi: [10.1258/jtt.2008.080412](https://doi.org/10.1258/jtt.2008.080412)] [Medline: [19246607](https://pubmed.ncbi.nlm.nih.gov/19246607/)]
31. Carroll AE, DiMeglio LA, Stein S, Marrero DG. Contracting and monitoring relationships for adolescents with type 1 diabetes: a pilot study. *Diabetes Technol Ther* 2011 May;13(5):543-549. [doi: [10.1089/dia.2010.0181](https://doi.org/10.1089/dia.2010.0181)] [Medline: [21406011](https://pubmed.ncbi.nlm.nih.gov/21406011/)]

Abbreviations

HbA1c: glycated hemoglobin

Edited by H Potts; submitted 17.01.12; peer-reviewed by K Clauson, J Kvedar, A Douglass-Bonner; comments to author 07.02.12; revised version received 04.04.12; accepted 23.04.12; published 08.05.12.

Please cite as:

Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR

Design of an mHealth App for the Self-management of Adolescent Type 1 Diabetes: A Pilot Study

J Med Internet Res 2012;14(3):e70

URL: <http://www.jmir.org/2012/3/e70/>

doi: [10.2196/jmir.2058](https://doi.org/10.2196/jmir.2058)

PMID: [22564332](https://pubmed.ncbi.nlm.nih.gov/22564332/)

©Joseph A Cafazzo, Mark Casselman, Nathaniel Hamming, Debra K Katzman, Mark R Palmert. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 08.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org>, as well as this copyright and license information must be included.

Original Paper

There's an App for That: Content Analysis of Paid Health and Fitness Apps

Joshua H West¹, PhD, MPH; P. Cougar Hall¹, PhD; Carl L Hanson¹, PhD; Michael D Barnes¹, PhD; Christophe Giraud-Carrier², PhD; James Barrett¹, BS

¹Computational Health Science Research Group, Department of Health Science, Brigham Young University, Provo, UT, United States

²Computational Health Science Research Group, Department of Computer Science, Brigham Young University, Provo, UT, United States

Corresponding Author:

Joshua H West, PhD, MPH

Computational Health Science Research Group

Department of Health Science

Brigham Young University

221 Richards Building

Provo, UT, 84602

United States

Phone: 1 801 422 3444

Fax: 1 801 422 0273

Email: josh.west@byu.edu

Abstract

Background: The introduction of Apple's iPhone provided a platform for developers to design third-party apps, which greatly expanded the functionality and utility of mobile devices for public health.

Objective: This study provides an overview of the developers' written descriptions of health and fitness apps and appraises each app's potential for influencing behavior change.

Methods: Data for this study came from a content analysis of health and fitness app descriptions available on iTunes during February 2011. The Health Education Curriculum Analysis Tool (HECAT) and the Precede-Proceed Model (PPM) were used as frameworks to guide the coding of 3336 paid apps.

Results: Compared to apps with a cost less than US \$0.99, apps exceeding US \$0.99 were more likely to be scored as intending to promote health or prevent disease (92.55%, 1925/3336 vs 83.59%, 1411/3336; $P<.001$), to be credible or trustworthy (91.11%, 1895/3336 vs 86.14%, 1454/3349; $P<.001$), and more likely to be used personally or recommended to a health care client (72.93%, 1517/2644 vs 66.77%, 1127/2644; $P<.001$). Apps related to healthy eating, physical activity, and personal health and wellness were more common than apps for substance abuse, mental and emotional health, violence prevention and safety, and sexual and reproductive health. Reinforcing apps were less common than predisposing and enabling apps. Only 1.86% (62/3336) of apps included all 3 factors (ie, predisposing, enabling, and reinforcing).

Conclusions: Development efforts could target public health behaviors for which few apps currently exist. Furthermore, practitioners should be cautious when promoting the use of apps as it appears most provide health-related information (predisposing) or make attempts at enabling behavior, with almost none including all theoretical factors recommended for behavior change.

(*J Med Internet Res* 2012;14(3):e72) doi:[10.2196/jmir.1977](https://doi.org/10.2196/jmir.1977)

KEYWORDS

mHealth; iPhone; app

Introduction

The use of mobile devices in supporting health behavior change is promising. Aside from expanded opportunities for users to access health information, mobile devices are becoming useful for facilitating the ongoing collection of personal data and

cueing behavior change at the most appropriate times [1]. One of the earliest applications was in 1998 when wristwatches were used to cue recovering addicts to complete logs regarding their behavior and feelings [2]. The most recent health applications have focused on a simple message system [3], mostly related to diabetes management [4-6] and smoking cessation [4,7].

Clinical care applications are most prevalent with fewer targeting preventive health behaviors [4].

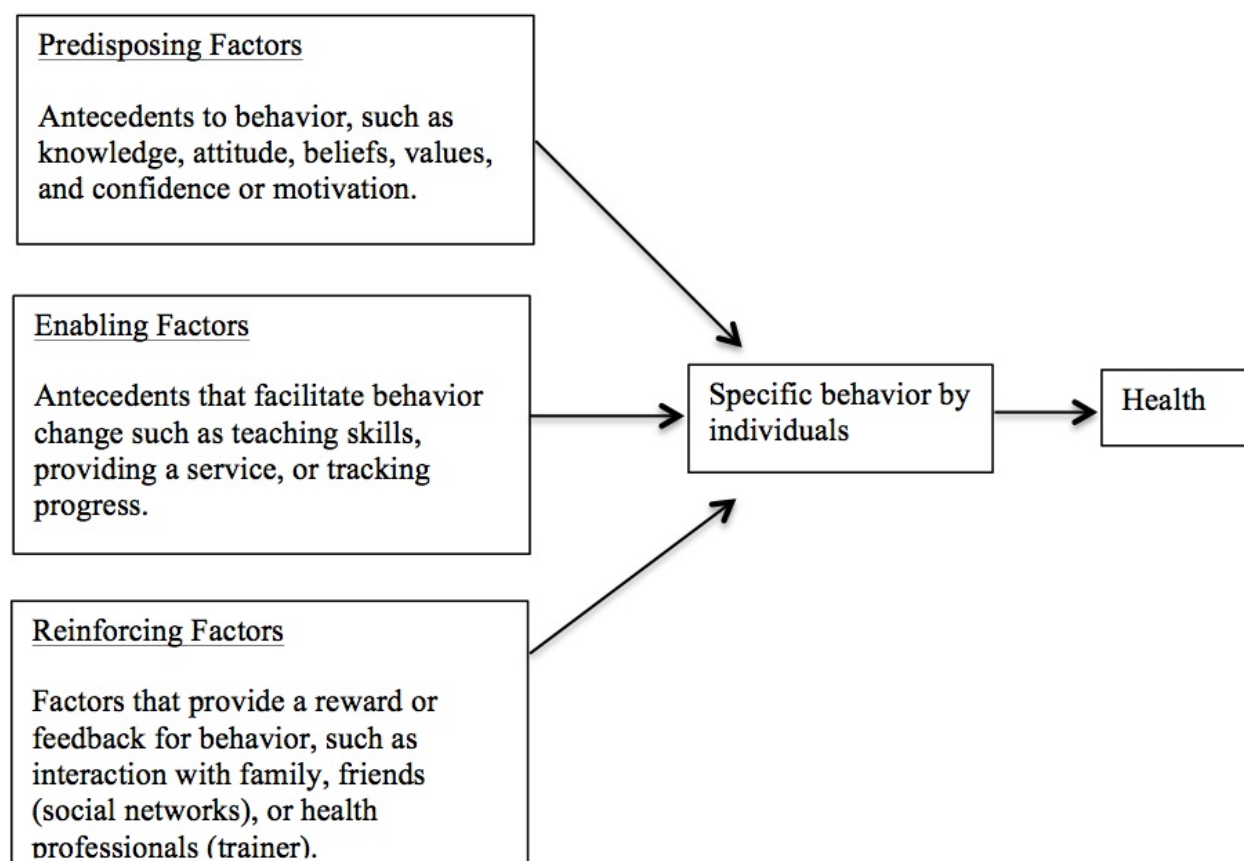
The recent emergence of smartphones has greatly enlarged both the reach and realm of possibilities for health purposes. In particular, it has provided a platform for developers to design third-party applications (apps), which expand the functionality and utility of these mobile devices. Apps are software programs designed specifically to run on mobile devices. In the “health and fitness” category in Apple’s App Store, developers have created thousands of downloadable apps for Apple’s mobile devices, which include the iPhone, the iPad, and the iPod touch. Since the launch of the Apple App Store in the US in July of 2008, more than 500,000 apps have become available with nearly 25 billion downloads [8]. By 2016, it is estimated that more than 44 billion apps will have been downloaded—which is equivalent to 6 app downloads for every person in the world [9].

There is growing interest and inquiry in understanding how mobile devices may influence health behavior [10]. Fogg introduced the idea of the functional triad [11]. The triad is a framework that delineates the role of devices in the human–device interaction. According to the triad, devices can be tools, mediums, or social actors. This triad is similar to widely accepted constructs pertaining to the Precede-Proceed Model (PPM) [12], namely predisposing, enabling, and reinforcing factors (Figure 1). Tools, which are similar to predisposing factors, increase the user’s capability. Mediums, which are similar to enabling factors, facilitate an authentic experience for users. Lastly, social actors, akin to reinforcing factors, assist the user in establishing and strengthening

relationships. For example, mobile devices become tools or predisposing factors when used to diffuse health information. Similarly, these devices serve as mediums or enabling factors when used by an individual to collect data regarding one’s personal health behavior. Apps can be considered social actors or reinforcing factors because they allow users to interact with social support networks or resources.

Physicians and other health care professionals have begun using many of these apps in their practices, primarily as reference tools [13,14]. As these technologies become more accessible to patients, practitioners may wish to recommend one of the many health and fitness apps to their clients [15,16]. There is currently little analysis in the scholarly literature about the quality of these apps [17]. The purpose of this study was to conduct a content analysis of app developers’ written descriptions of health and fitness apps, which interface with the iPhone. Recognizing that an analysis of these written descriptions falls short of a study design that might entail purchasing, downloading, and using the thousands of health and fitness apps available, the scope of this study aimed to analyze the same information that is available to consumers when considering an app purchase. Many companies develop apps for a variety of mobile devices, including makers of the Android operating system, which now sells more devices than Apple’s iPhone [18]. iPhone apps were chosen for this study because Apple’s App Store still provides the largest selection of apps for download and iPhone users download apps at about twice the rate of Android users, Apple’s closest competitor in the app marketplace [18]. Additionally, this study sought to analyze the written description for each health and fitness app to appraise its potential for influencing behavior change.

Figure 1. Causal relationship of Precede-Proceed Model factors influencing behavior and health, adapted from the Precede-Proceed Model for health promotion planning and evaluation in Green and Kreuter [12]; apps could address any one of the factors, or in some cases, address multiple factors.



Methods

Design

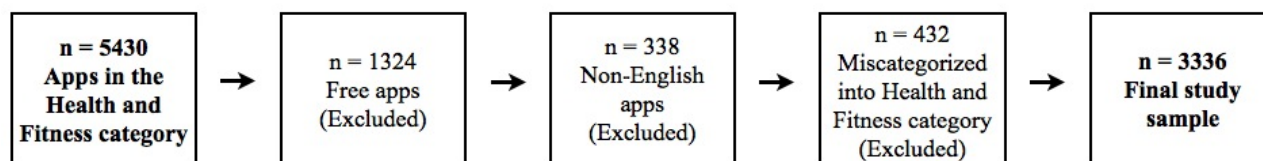
This study involved a qualitative content analysis of written descriptions of apps, which were provided by the apps' developers and were accessed through Apple's App Store via iTunes. iPhone users can view the App Store using the factory-loaded app, "App Store." The Health Education Curriculum Analysis Tool (HECAT) [19] and the PPM [12] were both used to guide the coding of the apps. The study authors were uniquely qualified to design and conduct this study. The authors had a combined 9 years of experience using iPhone apps for personal use, and 7 years of combined experience with apps in a university setting (eg, conducting research about apps and teaching computer science courses that focus on apps). One of the study authors served on a national planning committee focused on training faculty at institutes of higher education in implementing the HECAT, and three other study authors were trained in health behavior change theory and were very familiar with the PPM. The authors designed the coding methodology and instrumentation following a preliminary review of several hundred health and fitness apps.

The study authors trained two research assistants in the HECAT content areas, the PPM, and study mechanics (eg, data entry and coding the apps). Training sessions were held biweekly over the course of 4 weeks and each session lasted 1 hour. The research assistants were public health undergraduate students

in their final year of coursework. Both had completed classes that covered the HECAT content areas and health behavior change theories and models, which included a discussion of the PPM.

Sample

Apps included in this study sample came from the "health and fitness" category in the US Apple App Store (N = 5430), where descriptions of apps are easily accessed and viewed (Figure 2). The study sample was limited to English language apps within the health and fitness category; therefore, 338 apps were excluded. Freemium apps are a marketing strategy that have an option for free download, offer limited functionality, and then can be upgraded at a cost to the fully functional app. As of September 2011, freemium apps accounted for 72% of app downloads on Apple's App Store [20]. To avoid the duplication and subsequent coding errors that would arise from coding both the partially and fully functional versions of the same app, this study sample consisted of only paid apps; 1324 freemium apps were excluded. Based upon these inclusion/exclusion criteria, there were 3768 apps available for coding. Of these apps, 432 were miscategorized or misplaced under the health and fitness category (eg, alcohol buyers guides and restaurant locators) and were excluded. The final study sample (n = 3336) was comprised of only those apps that promoted health or prevented disease/injury. Furthermore, 73% of iPhone apps are paid [21]. Paid apps that met the inclusion/exclusion criteria for this study comprised a similar 76% of the total number of apps in the health and fitness category.

Figure 2. Selection of the study sample.

Measurement

Each app was coded for basic descriptive information, such as the app's name and price. Based upon the preliminary review of apps during study design, the authors concluded that many apps were either not intended to promote health or prevent disease, made claims that were not credible or trustworthy, or could not be recommended to a client for use in a professional setting. Coders assessed each app to determine if, from the perspective of a public health professional, the app was: (1) intended to promote health or prevent disease, (2) credible or trustworthy, and (3) generally recommendable to a client for use to improve health or prevent disease. Examples of apps that were coded as not intended to promote health or prevent disease included such things as restaurant locators or alcohol buyer's guides. Examples of apps that were coded as not credible or not trustworthy included such things as apps that claimed to predict the gender of your future baby, apps that predict the date of your death, and apps that purport being able to access extrasensory potentials of the mind. Examples of apps that were coded as not being recommendable to clients to improve health or prevent disease included such things as apps for remembering gym locker combinations or tips for applying makeup.

Next, each app was coded and categorized according to the health-related behavior it aimed to influence. Standard health education content areas included in the HECAT were used to facilitate this coding. These content areas were selected because they represent core areas of health education and behavior change, which comprise complete wellness as defined by the Centers for Disease Control and Prevention [19]. The HECAT's content areas, and subsequent coding categories for this study, included the following: alcohol, tobacco and other drugs, healthy eating, mental and emotional health, personal health and wellness, physical activity, safety, sexual health, and violence prevention. The authors reviewed hundreds of apps during the study design phase in order to inform the creation of the subcategories within each of the core HECAT content areas. After reviewing each app's description, the coders determined which content areas, and subcategories, were most descriptive of the behavior(s) addressed by the app.

Next, the PPM was used to code each app according to its level of anticipated influence to potentially change behavior. The PPM is a widely accepted health education framework [12] and was used in this study to guide the coding of the apps on several important dimensions related to behavior change. The PPM groups health behavior determinants into 3 main categories: predisposing factors, enabling factors, and reinforcing factors. For this study, predisposing apps were those utilities likely to precede behavior and which were cognitive- or affective-based. Predisposing apps were related to: knowledge or awareness of conditions or outcomes (eg, an app that provides cancer

statistics); providing information (eg, an app that presents information regarding ways to prevent adverse health outcomes); beliefs, values, or attitudes (eg, an app that discusses common reasons to avoid tobacco in an effort to assist the user in quitting smoking); and confidence or motivation (eg, an app that tries to convince you that you can change your diet). Coders coded apps as enabling if they were intended to be utilized, or occurred, at or around the same time as the desired behavior and if they facilitated behavior through teaching a skill (eg, an app with pictures and instructions on healthy stretching), providing a service (eg, an app that geo-locates places for physical activity), or tracking progress/recording behavior (eg, calorie counter apps). Reinforcing factors are the rewards received and the feedback the learner receives from others following adoption of a behavior, which may encourage or discourage continuation of the behavior. Coders coded apps as reinforcing if they interfaced with a social networking site (eg, apps with automatic upload to Facebook), provided encouragement from trainers/coaches (eg, an app that featured easy communication with a coach or trainer), and included an evaluation based upon the user's self-monitoring (eg, an app that provided automated feedback about user's reports of his/her physical activity).

Analysis

Inter-rater reliability was computed between the 2 coders on a subsample of 10% of the final dataset. The number of agreements were divided by the number of disagreements and the resulting level of concordance was 92%, which is comparable to levels used in previous content analysis research [22]. Disagreements between coders were discussed as a research team until a consensus was reached.

Apps were compared according to price to determine if the basic quality of more expensive apps increased, when compared to apps that cost less than US \$1. The price point of US \$1 was chosen for comparison because nearly half of the sample (42.3%) of apps was priced below US \$1. Frequencies of apps were computed according to each of the HECAT content areas. Additionally, the frequencies of apps were computed according to each of the dimensions of the PPM.

Results

Most of the apps in this study cost \$0.99 (1411/3336, 42.30%), 23.77% cost \$1.99 (793/3336), and 33.93% cost \$2.99 or more (1229/3336). Table 1 shows the differences between apps according to their listed purchase price. Apps that cost > US \$1 (\$0.99) were significantly more likely to be coded as intending to promote health or prevent disease ($P < .001$). Higher priced apps were also more likely to be coded as credible or trustworthy at being able to promote health or prevent disease ($P < .001$) and they were more likely to be coded as recommendable to a

client for the purpose of improving health or preventing disease ($P < .001$).

Each HECAT content area and the respective representation among study sample apps are included in Table 2. For this study, the HECAT content areas were not mutually exclusive. Indeed, many apps were coded under multiple content areas. Apps related to physical activity and personal wellness were the most common. Among physical activity apps, workout programs and workout monitors accounted for nearly all the apps. Apps in the personal wellness category were comprised of those relating to complementary and alternative medicine, sleep, remedies, disease-specific information (eg, information about heart disease), oral care, and hygiene. Healthy eating and mental health apps combined to account for just over 30% (1065/3336, 31.92%) of apps. Nutritional content of specific food items,

calorie tracking or counting apps, and healthy diet apps were the most common under the healthy eating category. In the mental health category, stress management apps comprised the largest subgroup, followed by meditation guides, and remedies, therapies and self-help treatments. Apps related to sexual and reproductive health; alcohol, tobacco, and other drugs (ATOD); and injury prevention were the least common in this study sample. Pregnancy or fertility calendars were the most common type of sexual/reproductive health apps, followed by postnatal care, and then prenatal care. Nearly all ATOD apps were designed to assist the user in his/her own addictions, with 1.53% (2/129) including tips on how to provide support and assistance to an addict. Apps coded under the injury prevention category were mostly related to first aid and emergency preparedness. One-third of injury prevention apps provided information about snakebites.

Table 1. Comparison of app credibility by price.

	n (%)			
	\$0.99	>\$0.99		<i>P</i>
This is an app intended to promote health or prevent disease	1411 (83.59)	1925 (92.55)	73.7	< .001
This app is credible or trustworthy	1454 (86.14)	1895 (91.11)	23.3	< .001
As a health care professional, I would use this app for my personal use or recommend it for use by one of my clients	1127 (66.77)	1517 (72.93)	16.9	< .001

Table 2. Frequencies of apps according to the Health Education Curriculum Analysis Tool (HECAT) content areas (N = 3336).

HECAT content area	n (%) ^a
Alcohol, tobacco, and other drugs	131 (3.93)
Help for addiction	118 (90.08)
Support for an addict	2 (1.53)
Other	17 (12.98)
Healthy eating	651 (19.51)
Calorie counters, journals, logs	213 (32.72)
Healthy recipes and cooking tips	114 (17.51)
Healthy diet-specific information	175 (26.88)
Nutritional breakdown of specific food items	236 (36.25)
Other	57 (8.76)
Mental and emotional health	414 (12.41)
Eating disorders	11 (2.06)
Stress management	255 (47.66)
Depression	41 (7.66)
Reference tests, diagnostic tools, information	18 (3.36)
Remedies, therapies, medication guides, self-help treatments	131 (24.49)
Meditation guides	160 (29.91)
Other	
Physical Activity	1108 (33.21)
Workouts, tips, ideas	696 (62.82)
Parks, facilities, directional maps	11 (0.99)
Race announcements, events	1 (0.09)
Monitors, measurement of workouts, logs, automatic recordings	496 (44.40)
Other	8 (0.72)
Violence prevention and safety	96 (2.88)
Attack alarms, notification noises	15 (15.62)
How-to guide for administering first aid	32 (33.33)
Information about snake bites	32 (33.33)
Emergency preparedness	25 (26.04)
Other	26 (27.08)
Personal health and wellness	962 (28.84)
Sleep	160 (16.63)
Oral care/hygiene	14 (1.46)
Disease/illness specific information	146 (15.18)
Remedies/medications/prescriptions	153 (15.90)
Goal setting	37 (3.85)
Beautification	28 (2.91)
Complementary and alternative medicine	310 (32.22)
Skin care	31 (3.22)
Other	156 (16.22)

HECAT content area	n (%) ^a
Sexual and reproductive health	243 (7.28)
Prenatal care	60 (24.69)
Pregnancy/fertility calendar	82 (33.74)
Postnatal care	72 (29.63)
Early parenting strategies and tips	16 (6.58)
Intimacy enhancer	25 (10.29)
Other	8 (3.29)

^a Apps could be coded in multiple categories.

Most of the apps were coded as either predisposing or enabling (Table 3). Only 6.65% (222/3336) of apps were reinforcing and 1.86% (62/3336) of apps included each PPM factor (ie, predisposing, enabling, and reinforcing). Predisposing aspects of apps mostly involved the provision of knowledge or attempts to raise awareness (713/1776, 40.15%). Enabling apps were most commonly coded as providing a service (1065/2181, 48.85%). Reinforcing apps were coded as such mostly because

they interfaced with social networking sites (eg, automatic uploads to Facebook), potentially facilitating feedback from peers regarding behavior. Over 40% (92/222, 41.44%) of reinforcing apps provided some type of feedback—sometimes automated and sometimes human-to-human communication—regarding the user's self-monitoring of his/her behavior.

Table 3. Theoretical classification of apps (N = 3336).

PPM Level	n (%) ^a
Predisposing	1776 (53.24)
Knowledge or awareness	713 (40.15)
Informative	1372 (77.25)
Beliefs, values, attitudes	378 (21.28)
Confidence or motivation	369 (20.78)
Enabling	2181 (65.38)
Teach a skill	810 (37.16)
Provide a service or sell something	1065 (48.85)
Track/record behavior	859 (39.40)
Reinforcing	222 (6.65)
Interfacing with social networking sites for encouragement	101 (45.50)
Encouragement, trainer support, coach	50 (22.52)
Evaluation based upon self-monitoring	92 (41.44)
All	62 (1.86)

^a Apps could be coded in multiple categories.

Discussion

The more expensive apps in this study sample were identified as more credible or trustworthy, more recommendable to clients in a professional setting, and more likely designed to promote health and prevent disease. These findings are consistent with Abroms et al [23] who reported that most of the apps scoring highest in adherence to established clinical practice guidelines for tobacco cessation were paid apps, thus establishing a relationship between quality and price. Together, these findings may indicate that apps with many functions simply cost more. While Heron and Smith [24] reported evidence of mobile

devices' efficacy for bringing about behavior change, the current study supports a recognition that such devices will require a level of sophistication likely to increase cost [25].

Personal health and wellness, physical activity, and (to a lesser extent) healthy eating apps were by far the most represented categories of apps. Each of these categories represent domains in which individuals historically feel empowered to bring about their own change. This may also reflect a perception that solutions to these problems are less complex. Furthermore, the worldwide rate of obesity has more than doubled since 1980 [26], which has perhaps increased the popularity of exercise and nutrition apps that claim to be self-helps for individuals

wishing to lose weight. The observation that ATOD apps were less common is of note considering the continued and considerable substance abuse problems in the United States [27-29]. One explanation for this, as Abroms et al [23] point out in their study of smoking cessation apps, is that app users may represent an elite group of affluent individuals, who may be less likely to suffer from substance abuse. As the diffusion of mobile devices increases to include a wider socioeconomic demographic, there may be an increased need for development of efficacious ATOD apps.

More than half of the apps in the current study were established upon predisposing factors, which are primarily knowledge-based. If most health and fitness apps available are simply predisposing factors, professionals should consider the added utility of these apps above and beyond traditional approaches (eg, self-help guides and manuals and reference books). Upon first inspection, apps in the current study appear more affordable (ie, US \$0.99) than traditional approaches and are consistent with the price of apps studied in other disciplines [30]. However, considering that all users access the App Store via a device (eg, iPhone, iPod, or laptop) that contracts with a data service provider for a monthly fee, the actual price to the user may indeed be higher, except in cases where users connect using a borrowed connection (eg, work or free Wi-Fi). However, these higher costs may be offset by the convenience associated with being able to consolidate functions into a single device. For example, being able to read about risk factors for heart disease and make personal calls all with the same device may be very appealing to many users. Nevertheless, despite the convenience of modern mobile devices, the health and fitness apps included in this study do not extend beyond what may be accomplished through traditional approaches that employ predisposing factors. Furthermore, given that many of these apps focus on the provision of information, future research should consider the medical accuracy of their content.

The most commonly coded theoretical classifications of apps in the current study were those based upon enabling factors, such as teaching skills, tracking progress, or recording actual behavior. This finding is supported by a recent systematic review of mobile device applications for diabetes management, in which Chomutare et al [6] report that tracking progress, an enabling approach, was among the most common features of apps designed for diabetics. Ravert et al [31] note that mobile devices may be used as a means to measure behavior in a convenient and immediate manner, which has historically been dependent on the users recalling and manually inputting data. Indeed, manual uploads to mobile devices regarding one's behavior may be preferred to more traditional recording practices, namely

pen and paper [32]. Additionally, more automated processes introduce the possibility of harnessing the devices' capabilities for monitoring, which alone may provide incentive to change behavior by allowing the user to monitor and report about their behavior change progress [33,34]. Figure 3 presents actual examples of study apps that represent predisposing, enabling, and reinforcing factors.

Few apps in the current study were found to include reinforcing factors, which are characterized by the provision of encouragement, evaluation, and the opportunity to interact with others. This finding is consistent with that of Abroms et al [23] and Chomutare et al [6] who reported that few apps connect the user to outside sources, including social support systems. This appears to be a missed opportunity given the capacity of emerging mobile device technology. Adapting from the PPM, the hallmark of reinforcing is the extent to which the apps connect the user with external systems or communities, such as social networking sites. As it relates to mobile devices, Patrick et al [1] and Heron and Smith [24] have referred to this process as ecological momentary interventions, or as Intille et al [35] call it, "just-in-time." Such interventions refer to apps that adapt as a result of data obtained from the user. These uses might be thought of as real-time behavior change support, where users can receive reinforcement via a pre-programmed virtual coach or actual human interaction. According to the PPM, these apps would represent a complete approach to changing behavior and should be most efficacious, which could be determined through research using randomized controlled trials.

Fjeldsoe et al [4] recognized the need for interventions utilizing mobile devices to be based upon theoretical principles. Public health professionals could partner with app developers to create apps that align with established behavior change theories [36], including those that would emphasize the reinforcing paradigm prominent in the PPM. Based upon the current study, the majority of existing apps are limited in their inclusion of reinforcing factors, which are considered vital in facilitating behavior change. Theory is critical in public health interventions and research because it aids in understanding how and why individuals, groups, and organizations behave and change [37]. Public health should take a more active role in the creation of apps, which include theoretical principles. Indeed, previous research has shown that mobile device interventions are most effective when they are based upon theory [3,10]. These interventions will become increasingly more available as the technology's penetration deepens [38] and could be a powerful tool for the public health professional in broadening their influence and in reaching previously isolated segments of the community [39].

Figure 3. A. iHome Remedies (factor-predisposing); B. The Benefits of Vitamins (factor-predisposing); C. Calorie Track (factor-enabling); D. iWeight Track (factor-enabling); E. DiabetesPlus (factor-reinforcing); F. ithlete (factor-reinforcing).



Limitations

Whereas this study represents one of the first of its kind, there are several limitations that should be considered when interpreting the results. First, the coders used the developers' own descriptions of their apps in order to code them. It is possible that some developers either overstated or understated the capabilities of their apps, which would have resulted in a misclassification for the purposes of this study. Actually downloading and using the apps was beyond the scope of this study and would not have been feasible for a sample of 3336 apps. Furthermore, the written descriptions of the apps comprise the same information that customers have available to them when they consider making a purchase. Second, the apps represented in this study may not include all of the apps potentially relevant to health and fitness. Just as the coders for this study identified apps that were not intended to promote health or prevent disease, it is possible that other relevant apps were equally misclassified into other App Store categories, which would have precluded their inclusion in this study. In addition, this study focused only on apps associated with Apple's App Store and excluded apps associated with other platforms (eg, Android apps). Whereas future studies may explore apps designed for other platforms, this study's focus

on iPhone apps was because this platform has the greatest number of apps available for download [40]. Lastly, the PPM and HECAT content areas guided the coding of the apps and appeared adequate to the researchers, but it is unclear to what extent these coding categories covered all of the types of the apps listed under this section in the App Store.

Conclusions

There are many apps available to those desiring to promote health or prevent disease. Nevertheless, practitioners wishing to recommend the use of third-party apps for such devices as Apple's iPhone should do so with discretion. It is recommended that practitioners be prudent when promoting the use of apps so as not to overstate their potential effectiveness. Based upon the study apps' descriptions, it appears that most provide health-related information or make attempts at enabling behavior, with almost none including all factors of the PPM recommended for behavior change. Furthermore, development efforts could target important public health behaviors for which few apps currently exist, such as substance abuse. Future research should extend recent work by Abroms et al [23] and actually test the efficacy of a large number of apps because little is known about their potential health utility.

Acknowledgments

This research was conducted through internal funds provided by Brigham Young University.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Apps coding form.

[[PDF File \(Adobe PDF File\), 42KB - jmir_v14i3e72_app1.pdf](#)]

References

1. Patrick K, Intille SS, Zabinski MF. An ecological framework for cancer communication: implications for research. *J Med Internet Res* 2005 Jul 1;7(3):e23 [[FREE Full text](#)] [doi: [10.2196/jmir.7.3.e23](#)] [Medline: [15998614](#)]
2. Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychol* 1998 Jan;17(1):48-52. [Medline: [9459069](#)]
3. Fjeldsoe BS, Miller YD, Marshall AL. MobileMums: a randomized controlled trial of an SMS-based physical activity intervention. *Ann Behav Med* 2010 May;39(2):101-111. [doi: [10.1007/s12160-010-9170-z](#)] [Medline: [20174902](#)]
4. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. *Am J Prev Med* 2009 Feb;36(2):165-173. [doi: [10.1016/j.amepre.2008.09.040](#)] [Medline: [19135907](#)]
5. Katz DL, Nordwall B. Novel interactive cell-phone technology for health enhancement. *J Diabetes Sci Technol* 2008 Jan;2(1):147-153 [[FREE Full text](#)] [Medline: [19885191](#)]
6. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. *J Med Internet Res* 2011;13(3):e65 [[FREE Full text](#)] [doi: [10.2196/jmir.1874](#)] [Medline: [21979293](#)]
7. Obermayer JL, Riley WT, Asif O, Jean-Mary J. College smoking-cessation using cell phone text messaging. *J Am Coll Health* 2004;53(2):71-78. [doi: [10.3200/JACH.53.2.71-78](#)] [Medline: [15495883](#)]
8. Miles S. Pocket-lint. Apple: Nearly 25 billion apps downloaded URL: <http://www.pocket-lint.com/news/44485/apple-25bn-app-downloads-worldwide> [accessed 2012-02-29] [[WebCite Cache ID 1330494839325301](#)]
9. CTIA-International Association for the Wireless Telecommunications Industry. 2011. 50 wireless quick facts URL: <http://www.ctia.org/advocacy/research/index.cfm/AID/12072> [accessed 2011-10-27] [[WebCite Cache ID 62kZxcSVZ](#)]
10. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med* 2011 Mar 1;1(1):53-71. [doi: [10.1007/s13142-011-0021-7](#)] [Medline: [21796270](#)]
11. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. In: Persuasive Technology: Using Computers to Change What We Think and Do. Waltham, MA: Morgan Kaufmann; 2002.
12. Green LW, Kreuter MW. Health promotion planning: An educational and environmental approach. In: Health promotion planning: an educational and environmental approach. Mountain View, CA: Mayfield Pub. Co.; 1991.
13. Oehler RL, Smith K, Toney JF. Infectious diseases resources for the iPhone. *Clin Infect Dis* 2010 May 1;50(9):1268-1274 [[FREE Full text](#)] [doi: [10.1086/651602](#)] [Medline: [20233061](#)]
14. Franko OI. Smartphone apps for orthopaedic surgeons. *Clin Orthop Relat Res* 2011 Jul;469(7):2042-2048. [doi: [10.1007/s11999-011-1904-0](#)] [Medline: [21547414](#)]
15. Wardrop, M. The Telegraph. Doctors told to prescribe smartphone apps to patients URL: <http://www.telegraph.co.uk/health/healthnews/9097647/Doctors-told-to-prescribe-smartphone-apps-to-patients.html> [accessed 2012-04-23] [[WebCite Cache ID 678nEPPZ](#)]
16. Department of Health. 2012. GPS to 'prescribe' apps for patients URL: <http://www.dh.gov.uk/health/2012/02/apps/> [accessed 2012-04-23] [[WebCite Cache ID 678nLrnVU](#)]
17. Hasman L. An introduction to consumer health apps for the iPhone. *Journal of Consumer Health On the Internet* 2011;15(4):322-329. [doi: [10.1080/15398285.2011.623567](#)]
18. Emerson R. The Huffington Post. Android apps more popular than iOS apps, ABI research study finds URL: http://www.huffingtonpost.com/2011/10/25/android-apps-ios-abi-research-study_n_1030595.html [accessed 2012-02-29] [[WebCite Cache ID 1330495764344650](#)]
19. US Department of Health and Human Services. Health Education Curriculum Analysis Tool. Atlanta, GA: Centers for Disease Control and Prevention; 2008.
20. Sutherland E. Cult of Mac. Freemium apps now account for 72 percent of downloads URL: <http://www.cultofmac.com/114733/freemium-apps-now-account-for-72-percent-of-downloads/> [accessed 2012-02-29] [[WebCite Cache ID 1330495954155386](#)]

21. Wauters, R. 28% of iPhone apps. Distimo: 57% of Android apps are free vs URL: <http://techcrunch.com/2010/07/05/distimo-june-2010/> [accessed 2012-04-23] [WebCite Cache ID 678nUfoWH]
22. Hall PC, West JH, McIntyre E. Female self-sexualization in MySpace.com personal profile photographs. *Sexuality & Culture* 2012;16(1):1-16. [doi: [10.1007/s12119-011-9095-0](https://doi.org/10.1007/s12119-011-9095-0)]
23. 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. [doi: [10.1016/j.amepre.2010.10.032](https://doi.org/10.1016/j.amepre.2010.10.032)] [Medline: [21335258](https://pubmed.ncbi.nlm.nih.gov/21335258/)]
24. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol* 2010 Feb;15(Pt 1):1-39. [doi: [10.1348/135910709X466063](https://doi.org/10.1348/135910709X466063)] [Medline: [19646331](https://pubmed.ncbi.nlm.nih.gov/19646331/)]
25. Kraft P, Yardley L. Current issues and new directions in Psychology and Health: What is the future of digital interventions for health behaviour change? *Psychol Health* 2009 Jul;24(6):615-618. [doi: [10.1080/08870440903068581](https://doi.org/10.1080/08870440903068581)] [Medline: [20205015](https://pubmed.ncbi.nlm.nih.gov/20205015/)]
26. World Health Organization. 2011. Obesity and overweight URL: <http://www.who.int/mediacentre/factsheets/fs311/es/index.html> [accessed 2012-02-29] [WebCite Cache ID 1330496491916691]
27. Centers for Disease Control and Prevention. 2010. Alcohol and public health URL: <http://www.cdc.gov/alcohol/> [accessed 2011-10-27] [WebCite Cache ID 62kaPcNR2]
28. Drug Addiction. 2010. Prescription drug addiction crosses all classes, demographics URL: <http://www.drugaddictiontreatment.com/types-of-addiction/prescription-drug-addiction/prescription-drug-addiction-crosses-all-classes-demographics/> [accessed 2011-10-27] [WebCite Cache ID 62kaTkZeJ]
29. Executive Office of the President of the United States of America. Epidemic: Responding to America's prescription drug abuse crisis. Washington, DC: White House; 2011.
30. Shuler C. iLearn: A Content Analysis of the iTunes App Store's Education Section. New York, NY: The Joan Ganz Cooney Center at Sesame Workshop; 2009. URL: http://blogs.edweek.org/edweek/DigitalEducation/iLearn_final%20formatted.pdf [accessed 2012-05-08] [WebCite Cache ID 67WS5Bwxj]
31. Ravert R, Calix S, Sullivan M. Research in brief: Using mobile phones to collect daily experience data from college undergraduates. *J Coll Stud Dev* 2010;51(3):343-352.
32. Boushey CJ, Kerr DA, Wright J, Lutes KD, Ebert DS, Delp EJ. Use of technology in children's dietary assessment. *Eur J Clin Nutr* 2009 Feb;63 Suppl 1:S50-S57. [doi: [10.1038/ejcn.2008.65](https://doi.org/10.1038/ejcn.2008.65)] [Medline: [19190645](https://pubmed.ncbi.nlm.nih.gov/19190645/)]
33. Wiehe SE, Carroll AE, Liu GC, Haberkorn KL, Hoch SC, Wilson JS, et al. Using GPS-enabled cell phones to track the travel patterns of adolescents. *Int J Health Geogr* 2008;7:22 [FREE Full text] [doi: [10.1186/1476-072X-7-22](https://doi.org/10.1186/1476-072X-7-22)] [Medline: [18495025](https://pubmed.ncbi.nlm.nih.gov/18495025/)]
34. Wiehe SE, Hoch SC, Liu GC, Carroll AE, Wilson JS, Fortenberry JD. Adolescent travel patterns: pilot data indicating distance from home varies by time of day and day of week. *J Adolesc Health* 2008 Apr;42(4):418-420. [doi: [10.1016/j.jadohealth.2007.09.018](https://doi.org/10.1016/j.jadohealth.2007.09.018)] [Medline: [18346668](https://pubmed.ncbi.nlm.nih.gov/18346668/)]
35. Intille SS, Kukla C, Farzanfar R, Bakr W. Just-in-time technology to encourage incremental, dietary behavior change. *AMIA Annu Symp Proc* 2003:874. [Medline: [14728379](https://pubmed.ncbi.nlm.nih.gov/14728379/)]
36. Chambers CD, Hughes S, Meltzer SB, Wahlgren D, Kassem N, Larson S, et al. Alcohol consumption among low-income pregnant Latinas. *Alcohol Clin Exp Res* 2005 Nov;29(11):2022-2028. [Medline: [16340460](https://pubmed.ncbi.nlm.nih.gov/16340460/)]
37. Rimer B, Glanz K. Theory at a Glance: A Guide for Health Promotion Practice. Bethesda, MD: National Cancer Institute, US Department of Health and Human Services; 2005. URL: <http://www.cancer.gov/cancertopics/cancerlibrary/theory.pdf> [accessed 2012-05-08] [WebCite Cache ID 67WSH7kWC]
38. Pew Internet & American Life Project. 2010 URL: http://www.pewinternet.org/~media/Files/Reports/2010/PIP_Mobile_Access_2010.pdf [accessed 2011-10-27] [WebCite Cache ID 62kaM3g0l]
39. Crilly JF, Keefe RH, Volpe F. Use of electronic technologies to promote community and personal health for individuals unconnected to health care systems. *Am J Public Health* 2011 Jul;101(7):1163-1167. [doi: [10.2105/AJPH.2010.300003](https://doi.org/10.2105/AJPH.2010.300003)] [Medline: [21566023](https://pubmed.ncbi.nlm.nih.gov/21566023/)]
40. Rosati J. Today's iPhone. Social Saturday: iOS apps vs android apps URL: <http://www.todaysiphone.com/2011/09/social-saturday-ios-apps-vs-android-apps/> [accessed 2011-10-27] [WebCite Cache ID 62kaEQYcf]

Abbreviations

ATOD: alcohol, tobacco, and other drugs

HECAT: Health Education Curriculum Analysis Tool

PPM: Precede-Proceed Model

Edited by G Eysenbach; submitted 27.10.11; peer-reviewed by T Miron-Shatz, O Ahmed, J Kim, L Fernandez-Luque, L Gualtieri, S Langrial, F Pappalardo, P Cipresso; comments to author 13.01.12; revised version received 23.04.12; accepted 25.04.12; published 14.05.12.

Please cite as:

West JH, Hall PC, Hanson CL, Barnes MD, Giraud-Carrier C, Barrett J
There's an App for That: Content Analysis of Paid Health and Fitness Apps
J Med Internet Res 2012;14(3):e72
URL: <http://www.jmir.org/2012/3/e72/>
doi: [10.2196/jmir.1977](https://doi.org/10.2196/jmir.1977)
PMID: [22584372](https://pubmed.ncbi.nlm.nih.gov/22584372/)

©Joshua H. West, P. Cougar Hall, Carl L. Hanson, Michael D. Barnes, Christophe Giraud-Carrier, James Barrett. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 14.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Researchers' Needs for Resource Discovery and Collaboration Tools: A Qualitative Investigation of Translational Scientists

Suresh K Bhavnani¹, PhD; Michael Warden², MBA; Kai Zheng^{3,4}, PhD; Mary Hill⁵, MS; Brian D Athey⁶, PhD

¹Institute for Translational Sciences, University of Texas Medical Branch, Galveston, TX, United States

²Elsevier, Inc., Ann Arbor, MI, United States

³School of Public Health, Department of Health Management and Policy, University of Michigan, Ann Arbor, MI, United States

⁴School of Information, University of Michigan, Ann Arbor, MI, United States

⁵Medical School Information Services, University of Michigan Medical School, Ann Arbor, MI, United States

⁶Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:

Suresh K Bhavnani, PhD

Institute for Translational Sciences

University of Texas Medical Branch

301 University Blvd

Galveston, TX, 77555-0331

United States

Phone: 1 (409) 772 1928

Fax: 1 (409) 772 1968

Email: skbhavnani@gmail.com

Abstract

Background: A critical aspect of clinical and translational science (CTS) is interdisciplinary and collaborative research, which increasingly requires a wide range of computational and human resources. However, few studies have systematically analyzed such resource needs of CTS researchers.

Objective: To improve our understanding of CTS researchers' needs for computational and human resources in order to build useful and useable supporting informatics tools.

Methods: We conducted semistructured interviews of 30 CTS researchers from the University of Michigan, followed by qualitative analysis of the interview transcripts.

Results: The analysis identified three recurring themes: the need for the federation of information, the need to address information overload, and the need to humanize computing, including strong and well-informed views about the use of social networking tools for research collaboration. These findings helped us to narrow down the available design choices for assisting CTS researchers, and helped to identify potential deficiencies of well-known theoretical frameworks used to guide our study, with suggestions for future remedies.

Conclusions: The user needs identified through the study, along with concrete design suggestions, provided key design, methodological, and theoretical insights, which are being used to guide the design and development of a CTS resource portal. The results and interview instrument should be useful to other institutions with Clinical and Translational Science Awards that face similar challenges related to helping CTS researchers make more effective use of computational and human resources.

(*J Med Internet Res* 2012;14(3):e75) doi:[10.2196/jmir.1905](https://doi.org/10.2196/jmir.1905)

KEYWORDS

Translational research; collaborative computing; Clinical and Translational Science Awards; CTSA; Resource Discovery

Introduction

Despite billions of dollars invested in biomedical research over the past five decades, there is a growing realization that our

ability to generate medical breakthroughs far exceeds our ability to apply those results in improving clinical care and population health. The US National Institutes of Health responded to these challenges by advocating and funding a new approach known

as clinical and translational science (CTS) [1]. The goal of CTS is to “accelerate the translation of laboratory discoveries into treatments for patients, to engage communities in clinical research efforts, and to train a new generation of clinical and translational researchers” [2].

The overall CTS approach is intended to encourage the 2-way translation of scientific discoveries between bench science, clinical research, and community-engagement activities, with the ultimate goal of improving human health. The growing acceptance of this approach has increased the motivation for researchers to work with collaborators across multiple disciplines. This shift has resulted in the need to use a wider range of computational and human resources. For example, basic scientists who wish to translate results from the laboratory to the design of clinical trials need access to computational resources within and across disciplines such as gene expression analysis and patient registries, and to identify and work with collaborators such as data mining experts and clinicians.

In recognition of these needs, several recipients of the Clinical and Translational Science Awards (CTSA) such as Harvard University and Vanderbilt University have already invested substantial effort in building Web portals to facilitate the discovery and use of computational and human resources. Furthermore, almost US \$30 million in federal funding has been allocated to CTSA sites for creating tools that enable researchers within and across institutions to communicate, collaborate, and discover resources more efficiently and effectively [3]. Other recipients of CTSA have partnered with commercial social networking platforms to create new resources for CTS researchers [4]. However, while some publications describing these efforts have targeted specific aspects of building research networking systems such as their interoperability [5], few studies have systematically analyzed the resource needs of CTS researchers to inform the design of such systems. For example, a recent set of studies (eg, [6-8]) conducted at the University of Pittsburgh attempted to identify the general services and functions needed specifically for locating experts to facilitate collaboration, leading to the design of a system called DigitalVita. However, to the best of our knowledge, there appears to be no published study that has directly probed the broad computational and human resource needs of CTS researchers and engaged them in designing an informatics solution.

We were therefore motivated to conduct a systematic investigation to directly engage CTS researchers in articulating their resource needs and to help guide the design of future solutions. Systematically understanding users' needs and engaging them in the design process is important for two reasons. First, users can help narrow down the range of possible solutions and avoid arbitrary choices purely based on expert opinions rather than on facts and evidence. For example, many CTSA institutions are exploring whether to build social networking tools for researchers but lack guidelines to help make such a decision or to identify key functionalities to implement. Second, empirical studies have repeatedly shown that computerized systems have been poorly adopted and underused in health care institutions, which is often due to the mismatch between delivered designs and the preferences and

constraints of the intended user population [9]. Such low adoption rates are underscored by technology acceptance models (TAMs), which have shown that the user acceptance of a technology is strongly influenced by the perceived usefulness and perceived ease of using that technology (eg, [10]). Furthermore, prior research has shown that involving potential users early in the development process of systems can help increase ownership of the final product and hence the likelihood of adoption [11].

In this paper, we describe the results of a study in which we engaged junior and senior CTS researchers at the University of Michigan in a qualitative study. Our goal was to identify the computational and human resource needs of the researchers, elicit concrete suggestions for the design of future systems, and engage the researchers in an iterative participatory design process.

Methods

We conducted our study at the University of Michigan (a CTSA site), which has 7600 faculty members, of whom 750 have been served by the CSTA and therefore can be considered CTS researchers. The University is also home to information technology departments that develop and customize information technology solutions across the medical campus. The leadership at the University of Michigan strongly recommends faculty and students to engage in interdisciplinary collaborations. Furthermore, since its establishment, the CTSA at the University of Michigan has received an increasing number of requests for accessing a wide range of human and computational resources. Given this need, and because the resource needs of CTS researchers had never been systematically analyzed, the CTSA leadership strongly supported conducting such a study. We therefore designed a semistructured interview study inviting a wide range of CTS researchers to help address the research question: *What are the computational and human resource needs of CTS researchers, and how can we translate them into designing useful and useable informatics solutions supporting interdisciplinary CTS research?*

The goal of the semistructured interviews was to enable an understanding of the motivations for the resource needs at a conceptual level, which would then guide the design of a survey with the goal to elicit information about specific resources (eg, project management applications) required by the researchers. We were also motivated to build a participatory design [12] relationship with the researchers so that they could be involved in the iterative refinement of future informatics solutions.

The design of our study was guided by two existing theoretical frameworks. First was the TAM and its latest extension, referred to as the unified theory of acceptance and use of technology. This theory postulates that performance expectancy (perceived usefulness), effort expectancy (perceived ease of use), social influence, and facilitating conditions are four important factors underlying end users' decision to accept or reject a technology, in addition to other moderating factors such as age and experience or seniority [10,13,14]. Some researchers (eg, Venkatesh et al) argue that this family of models accounts for a considerable proportion of behavioral intention variance in

explaining individuals' technology acceptance and usage decisions [13]. Second was technology-mediated collaboration research [15-17], whose results include (1) the importance of face-to-face collaborations for building trust, and (2) the importance of designing new collaboration technologies that are compatible with existing norms and practices so as to increase the probability of adoption.

Study Design

We used semistructured interviews to collect research data. Semistructured interviews are a well-known qualitative method [18] used in a wide range of fields including human-computer interaction, sociology, and medical informatics. The method is most useful when topics of research interest have been identified, but there is a lack of understanding of those topics about which to ask structured questions such as in a survey. The goal of the semistructured interview method is, on one hand, to focus an interview based on an ordered list of predetermined questions and, on the other hand, to enable the interviewer to explore issues that emerge during the interview, often leading to unexpected insights. This is achieved by asking open-ended, predetermined questions to enable users to discuss a topic. Depending on the answers, open-ended questions are then often followed opportunistically by carefully worded nonleading prompts to encourage continued elaboration, and probes to explore emergent issues or to guide the discussion in promising directions. Such interviews typically amass a large amount of in-depth qualitative data, and the studies are generally conducted with tens rather than hundreds of participants.

The two theoretical frameworks discussed above helped to guide the overall design of the study, in addition to the semistructured interview instrument. The theories on individual technology acceptance motivated us to include both junior and senior researchers in our study, as they might possess different perspectives regarding CTS research and therefore distinct needs for resources. In addition, the theories prompted us to include in the interview protocol key questions soliciting important theoretical constructs, such as the researchers' perceived barriers to using the proposed informatics tools for conducting

interdisciplinary translational research, and the social contexts in which they are situated that may convey salient social cues influencing their own beliefs of, and attitudes toward, such tools. Furthermore, the research on technology-mediated collaboration motivated us to ask three broad questions that were specifically related to the research process: how the researchers conducted or supported research activities, the nature of their scientific collaborations, and the tools that they used to conduct such research.

By using guidelines for interview design [19], we operationalized the above theoretical constructs into an interview instrument that proceeded from demographic and background questions to descriptive questions and broader opinion questions, and ended with user appraisals of two hypothetical designs: a Web portal concept integrating scattered information sources, and a digital curriculum vitae system that aggregates information about faculty and facilitates expertise finding. The resulting instrument consisted of 10 questions divided into 4 sections: (1) introduction and consent, (2) background and role, (3) research process, and (4) suggestions for supporting translational research. The interview protocol is provided in [Multimedia Appendix 1](#).

To recruit the CTS researchers, we used a 2-step snowball sampling method [20]. First, we requested 15 researchers who held leadership positions in the CTSA at the University of Michigan to participate in the study. Then, we requested the above researchers to identify junior faculty whose work was closely related to CTS. Among the junior faculty identified, we selectively invited a subset to participate in the study in order to ensure the representativeness of major disciplines and academic departments. The overall recruitment method resulted in 15 senior and 15 junior CTS researchers. These researchers were affiliated with a broad range of schools (eg, medicine, public health, nursing, and pharmacy) and had different backgrounds (eg, basic science, clinical and health services research, and community-based research), and therefore represent the breadth of CTS as it is currently conceptualized. [Table 1](#) summarizes the characteristics of the study participants.

Table 1. Participant characteristics (N = 30).

Characteristic	Total number of participants	Junior investigators	
		N	%
School			
Medical School/Health Systems	23	12	52%
School of Public Health	4	2	50%
School of Nursing	2	1	50%
College of Pharmacy	1	0	0%
Gender			
Female	12	9	75%
Male	18	6	33%

Each of the interviews was conducted by an information technology business analyst (MW) and a biomedical informatics faculty member (SKB or KZ). All interviews were conducted in the office of the researcher. Each interview lasted

approximately 90 minutes and was digitally recorded in audio format. The University of Michigan Institutional Review Board reviewed and approved the research protocol.

Data Analysis

The 30 audio-recorded interviews were transcribed by a professional transcriber resulting in approximately 600 pages of text. The data were subsequently analyzed in 3 steps. First, two analysts (MW and KZ) used the technique of open coding and categorization to iteratively annotate sections of the transcripts. This was achieved by using the constant comparison method [21] facilitated by the NVivo qualitative data analysis software (QSR International, Doncaster, Victoria, Australia). Second, a third analyst (SKB) independently analyzed and refined the thematic coding results. Third, all three analysts used affinity diagrams to reach a consensus on the final categorization of the data and identified implications of the results for a concrete design proposal.

Results

Analysis of the qualitative data helped to identify three emergent interrelated themes (and their subthemes), which captured the computational and human resource needs of CTS researchers. We first present evidence for these themes, followed by an analysis of the nature and causes of sentiments underlying those needs.

Emergent Themes Related to Resource Needs

As Table 2 shows, analysis of the interviews helped to identify three interrelated themes: (1) the need for the federation of information, (2) the need to address information overload, and (3) the need to humanize computing. In the following sections, we discuss each of these categories with examples of verbatim quotes provided by the study participants.

Table 2. Themes of user needs and design suggestions identified through the qualitative analysis of the interview transcripts.

Theme	Subtheme	Description
1. Need for the federation of information		
	Discoverability	Difficulty in discovering new information and resources
	Structured vs scattered information presentation	Disorientation caused by scattered information and resources
	<i>Design:</i> centralized access to resources	Desire for a centralized, authoritative location for information and resources
2. Need to address information overload		
	Relevant vs irrelevant information	Concern that informatics solutions will cause more information overload
	Push vs pull of information	Concern about how to balance the push and pull of information
	<i>Design:</i> comprehensible and personalized information	Desire for information presentation that is comprehensible and filtered based on personal preferences
3. Need for humanized computing		
	Human vs computer aided	Negative perceptions of tools that assume that research expertise can be found without humans
	User control vs automatic control	Negative perceptions of tools that will result in loss of control
	<i>Design:</i> human in the loop	Desire for human-aided services (eg, “concierge services” and “online consultation”) to facilitate finding research expertise

The Need for the Federation of Information

A common theme in the data was the need for the federation of information in three overlapping subthemes. First, there was an awareness of the large number of scattered resources across the clinical, research, education, and administrative systems. For example:

there are so many different sites and so many different tools—like you don’t know where to start

Next, several stated that, while new tools are beneficial to their research, they had difficulty *discovering* that they existed.

Our problem now is we’ve got [new resources and research expertise] and some of the people don’t even know we have [them]...somewhat the hardest part is figuring out what resources are out there, and who is out there

The above general awareness of useful but scattered and difficult-to-discover resources resulted in a strong desire for a central location where the resources could be accessed.

...the more things you can bring under one portal would be fantastic, and that sounds like something that would be very useful if it’s customizable

The results therefore confirmed that CTS researchers were indeed attempting to use a wide range of resources, but that those resources were highly scattered and it was difficult to know that they existed.

The Need to Address Information Overload

The researchers expressed concerns that any solution for integrating scattered resources needed to be designed carefully so that it would not further exacerbate the issue of information overload. They expressed this need in two ways. First, they expressed the need that system designers should develop systems that help to quickly distinguish between relevant and irrelevant information.

I know your business is really more of managing the data and mining it and stuff, but I mean people that are in the interface business, I mean we’re in

information overload...how do you distinguish between important information and nonimportant information? Because no one wants to waste their time.

In the absence of such tools, they had developed strategies to cope with information overload, such as using manual collaborative filtering techniques.

I skimmed through my emails quickly and deleted most of the event announcements without reading in full detail...I know that if something is really important to me, it will come back again later, for instance referred to me by my colleagues

Second, they expressed concerns about how to balance the push and pull of information. On one hand, the researchers were concerned that an excessive amount of information of little relevance will be pushed to them. On the other hand, if information is not pushed to everybody, the collaborative filtering strategy may no longer work.

You need like an optimal level of notification. If it's too much, people will disregard it. If it's not enough, you know...I don't think I'll look at another website just to see if potentially there's some collaborator.

The researchers provided several recommendations to address the information overload problem. These included interfaces that structured information for easy comprehension:

...so much information overload that you don't see the forest through the trees...[Design] something that doesn't have so much there that you don't in a sense know where to begin

Another suggestion was to build a system that provides tailored content based on each user's research area:

I am afraid I will not have time to customize it extensively, so I would much prefer if the tool can use some intelligence to predict my preferences based on my research areas.

The Need for Humanized Computing

The researchers expressed strong concerns about future tools that would dehumanize interpersonal relationships that are crucial for successful research collaborations. For example, they expressed suspicion toward the usefulness of the concept of automated research-expertise finding:

I realize there's a lot of informatics enthusiasm for collaboration ware for things like basically medical and professional versions of Facebook...That is beloved of informaticians, but frankly I don't think it's really going to fire up a lot of other people

In particular, there was widespread concern among junior as well as senior researchers that social networking applications would not be able to convey information about potential collaborators, such as whether a researcher is trustworthy. Such information was often implicit in personal referrals or direct communications. For example, a junior researcher stated

[Such tools] can find expertise; however, what it lacks is "personal touch"...if my boss says you should talk

to [name of a researcher]...like, he knows this person and he knows me...there is this personal touch because he is vouching for this person and there is some element of trust

Some of the senior researchers were concerned that being automatically enlisted in social networking tools would result in a large number of requests for new collaborations that they did not need. Such reservations, as one researcher noted, lead to the paradox of using social networking tools for research:

...the paradox is the types of people that you most want to have their information up-to-date because they might be able to be the most helpful to other people are the ones that are least like[ly] to do it...the junior faculty that have a lot of time will make sure that everything about themselves is amazingly accurate, but how many people are going to need to find a junior faculty member...?

The researchers provided several concrete recommendations for designing future systems that could overcome the concerns of dehumanized systems. These included systems that incorporated the human in the loop for finding resources:

...you basically say [on the website] "Did you get what you need, and if not type [it] in. If you could wait a couple of hours someone will respond by email," but at the bottom it's like "Don't leave this page unless you got what you wanted."

Another recommendation was for displaying social information of resource usage such as which tools are used by whom, with the goal of helping discover new resources based on what others are using.

Sentiments Underlying Resource Needs

While the above emergent themes directly addressed our goal of identifying the overall resource needs of CTS researchers, we were unprepared for the strength of the sentiments expressed during the interviews. During the interviews, the researchers required almost no prompting to discuss their needs, and most interviews ran over the 1.5 hours allotted, with 2 researchers inviting us back for follow-up interviews. We therefore probed deeper into the interview transcripts, in addition to reflecting on our personal face-to-face encounters with the researchers, to uncover cross-cutting themes that could explain the strong sentiments expressed during the discussions on resource needs.

Our analysis revealed that the strength of the sentiments did not appear to arise from an abstract conception or ignorance of new technologies, but rather were grounded in concrete interactions with university-based and contemporary Web-based technologies. For example, when discussing the need for a portal that federated resources, a researcher reflected on his use of a university-based system for institutional review board reviews:

When I do my reviews for the MCRU [Michigan Clinical Research Unit], the GCRC [General Clinical Research Center] legacy, I try to do them directly on eResearch. Some of my colleagues just threw up their hands and said: "Just send me a paper copy because I just can't do it. It's so inefficient that I don't even

want to learn. It's so arcane and so unwieldy I just can't even deal with it. So just send me a paper copy or I'll have my administrative assistant print off a paper copy, and I'll mark up the copy and send it back to you because I can't deal with it."

In the same vein, another researcher complained about the large overheads related to the CTSA-wide interactions:

This thing has grown into some kind of enormous effort that we didn't anticipate. It's kind of sad when you look at these proposals...oh, we will participate in the national level, blah, blah, and you say "Be careful what you wish for, guy"...it's an unfunded mandate basically...It's out of control. I think there's too much. There's too many wikis, too many things, too much need to interact with others, and this without it being paid for, and knowing this is a real problem.

In addition to experiences with university-based systems, the researchers were not naïve about contemporary Web-based systems, with equally strong sentiments. For example, a researcher, when discussing his use of browser Favorites, stated:

You read journal articles and you decide there's something new, and you might visit [it]. It references a website, and you might visit the website and decide "Well that's pretty good—I'm going to put it on my Favorites list." So I might put it on a Favorites list and say "Do I really go [to the Favorites list]? Do I gain anything by going there?"...I know I could do a Google search, and get [to the saved website] if I needed to.

Another researcher described the tensions involved in maintaining the accuracy of information in systems designed for enabling collaboration:

I mean there's a bit of a paradox here in that as you get busier, you're less likely to do these things and to check whether the information is accurate or whatever. There's so many different things that get sent to me about the U of M alumni club and stuff like that. I have never gone and looked to see if it's accurate. If it's accurate, it's accurate; [if] it's not, then people won't get a hold of me...So I think that we just have to be mindful of this as we construct any kind of [system].

However, the strongest sentiments pertained to the use of social networking tools (eg, Facebook) for research collaboration:

Teens are really cool on Facebook because they're big into the social networking aspects of it, but it's the social interaction that draws them there. They're there because their friends are there. If you're trying to build an online community where people who don't even know each other yet, you don't have that draw. It's always going to be a little in-club that's into that kind of thing. You're not going to get the people you're looking for. If you want the people you're looking for, face time, there's no substitute. Facebook is not a substitute for face time. You've got to get

people together in a room who wouldn't normally be together to hear about what they're talking about.

Besides being well informed, this sentiment about social networking did not seem related to a generational divide, based on our interview of a junior researcher aged 27:

I have a Facebook account but I never check it—just if people put photos on it—I use it to go look up the photos. But I'm not into that very much.

The sentiments surrounding resource needs therefore appeared to be grounded in real-world experience and knowledge of contemporary systems, a realization that strongly influenced our conceptions of how to design future solutions.

Discussion

Most biomedical applications are built without an adequate consideration of user needs and design involvement, which has been a major impediment to system adoption and long-term acceptance [9,22]. In most cases, users are asked for opinions after the systems are built or purchased, which strongly biases the developers toward asking questions and making convenient modifications that protect investments and prior design decisions, rather than serving the real needs of the users. When deployed systems fail to be adopted, it is often too late and difficult to make the significant changes needed for success, resulting in wasted time, effort, and resources [9,23]. This point is critical, as large amounts of funds are being spent to develop systems with perceived uses, rather than uses that are grounded in the needs of actual users.

Our approach was aimed to avoid such pitfalls by (1) probing the needs of users *before* system design and development began, (2) using the semistructured interview method, which allowed us to ask targeted questions while at the same time enabling users to express unexpected views that directly confronted our hypotheses and biases, (3) soliciting opinions for potential solutions, and (4) increasing the perception of ownership among a few key potential users, who could become champions to promote the informatics tools eventually developed.

Our main goals were to quickly narrow down the possible choices of system designs that would be useful and usable for CTS researchers, and to increase the chances for quick, widespread, and sustained adoption of the resulting system. To the best of our knowledge, this is the first study to systematically identify the resource needs of CTS researchers, with a focus on eliciting concrete design suggestions. Furthermore, our study was informed by the well-established and extensively validated TAM framework, in addition to extensive research in technology-mediated collaboration. Both frameworks guided us in the design of our study, enabling us to ask relevant questions during the semistructured interviews, which quickly yielded rich participant accounts that were critical to our research question. Furthermore, the 2-step snowball sampling method for selecting researchers from a wide range of backgrounds quickly enabled us to achieve those goals.

The resulting targeted study revealed that the researchers were frustrated about not being able to easily discover computational and human resources that were potentially useful, but were often

hidden and scattered. However, the enthusiasm for a centralized location for such resources was tempered by cautions about developing solutions that increased information overload and dehumanized computing.

While the above results might not appear novel in themselves, they helped us make important decisions about how to proceed with our developmental efforts. However, we were unprepared for the strength of the sentiments expressed during our discussions of resource needs. These sentiments did not appear to be naïve responses to technological solutions, but rather were based on real-world experience of university-based and contemporary Web-based tools. In particular, we were struck by the strong sentiments specifically against borrowing the generic *social networking* tools (eg, Facebook) for the purposes of research collaboration without a thoughtful redesign. Based on several debates within the design team, this finding has forced us to reexamine how current social networking tools (used essentially to maintain weak ties [24] and therefore requiring a low degree of trust) could evolve into a new class of research networking tools (used for finding and sustaining research collaboration, and therefore requiring a high degree of trust). After all, it is well known that current social networking tools are rarely used to find and establish new contacts [24]. Research networking tools should, therefore, be carefully designed to address the complex combination of trust and cultural issues required for scientific collaboration before they can be expected to be widely adopted for finding and establishing new research collaborations. Furthermore, we found senior researchers who possess the expertise and resources desired by junior faculty members to be least enthusiastic about using and contributing to such tools. Their lack of motivation to contribute to a research network could severely affect the network externality of such tools, hence diminishing their value.

The overall results have several similarities and differences with the studies conducted at the University of Pittsburgh [6-8]. Methodologically, although both of our approaches had the similar goal of engaging end users to determine their needs and to guide design, the Pittsburgh studies were different in that they (1) were motivated in part by the low adoption of an existing custom university system designed for helping faculty to establish collaborations [7], (2) were focused specifically on tools for research collaboration [8], and (3) recorded the content of the 27 interviews of researchers through handwritten notes after each interview was completed [6,7]. In contrast, our interviews did not focus on specifically understanding the use and design of collaboration tools, but rather were focused on situating our understanding of computational and human resource needs in the broader context of research work practices of translational scientists. Furthermore, the digital recordings provided verbatim accounts, which enabled us to recall, detect, analyze, and present the nuanced sentiments underlying the emergent themes. However, despite these methodological differences, there were several similarities in the results. Both sets of studies revealed that researchers often used existing real-world connections to determine new collaborations and had low motivation for updating their public online profiles. In fact, Schleyer and colleagues acknowledge that senior, more well-established researchers “are so well-informed and

well-connected that they, on average, will outperform any electronic system” [6]. They conclude by noting that research networking tools might therefore be more useful for junior researchers, contingent on an overall critical mass of researchers adopting the system. Future research should therefore investigate under which conditions junior researchers will use research networking tools, especially given the widespread availability and low cost of using powerful online tools such as Google, combined with the strong interpersonal advantages of using their mentors’ well-established collaboration networks.

The above findings have motivated us to pay careful attention to the sociotechnical issues of the researchers’ social status in the organization, and to provide appropriate incentive structures for collaboration and contribution when designing future solutions. As discussed below, these overall results have direct design, methodological, and theoretical implications.

Design Implications

The three emergent themes, and our subsequent understanding of the nature and causes of the strong sentiments underlying them, helped us to narrow down the available design choices for assisting CTS researchers. Because of the urgent need to federate scattered resources at the University of Michigan, we prioritized our efforts to first target that goal by building a CTS resource portal, keeping in mind that it could be scaled up to include future resources such as research networking. Accordingly, our design provides functionality that includes (1) a scalable system that federates access to a wide range of resources, (2) a dashboard view (possibly using the *portlet* technology [25]) of current projects and other information that is initially customized based on the researchers’ backgrounds, (3) the ability to modify the displayed resources through easy customization, (4) the ability to discover resources that are ranked based on a user’s profile and what other researchers are using, and (5) a concierge service to guide researchers to humans if they are unable to find specific computational or human resources. These functionalities directly address the emergent themes in our study.

To enable the researchers to continue to be engaged in the design process, a prototype of the above system was implemented and presented to a subset of the researchers for feedback and design input. Our preliminary analysis suggests that the researchers have responded positively to the prototype. Our future research includes development of a survey to solicit similar feedback about the prototype from a larger number of CTS researchers beyond those we originally interviewed, and iterative refinement of the system based on the continued participation of the intended users.

Theoretical and Methodological Implications

While our research focus was to precipitate a participatory design process, where we work closely with the researchers on progressive refinements of our informatics solutions, neither the technology acceptance theories nor the research on technology-mediated collaboration prepared us for the strong sentiments during our discussions on resource needs. These sentiments appeared to be directly related to past experiences with identical or similar technologies. While past experience

could be theoretically interpreted as a behavioral antecedent to existing constructs of the TAM (eg, to performance and effort expectancy), it manifested so distinctively in our study that we believe past experience deserves closer attention in its own right, both theoretically and practically. Therefore, we believe our results suggest that, in future research, additional theories that have explicitly modeled past experience, such as those developed by researchers in marketing science, could be incorporated into studies of the design and acceptance of information systems [26-29]. Accordingly, we recommend that designers of future systems for CTS researchers (a unique user population with the specific goal of translational research) pay close attention to their end users' past experiences with identical or similar technologies so that such experiences will be directly addressed in the design and presentation of new systems, with the goal of improving the chances of successful adoption and sustained use of the informatics solutions.

From a methodological perspective, it is also pertinent to note that over the course of our study and its presentations to various stakeholders, we have encountered strongly polarized views for the results. On one hand, translational researchers outside our study pool strongly identify with the results and often have their own contextualized accounts of similar sociotechnological issues. We have anecdotally encountered at other CTS sites almost identical issues related to computational resources that are not consolidated, difficult to use, and designed with strong assumptions of user needs, resulting in frustrations among researchers. On the other hand, technology developers and administrators, who are often suspicious of empirical studies, strongly criticize the results, using the argument that users are known to have difficulty articulating what they need and often base their needs narrowly on existing systems and biased views of new technologies. Instead, developers and administrators often propose a strategy to "build the system and users will come." Unfortunately, such a strategy has not had a very good track record of sustained success, especially in the biomedical

domain [9], often resulting in mandated use of poorly designed systems, coupled with contrived statistics for their success.

As professionals trained in the study of human-computer interaction and sociotechnical issues, we are acutely aware of such issues and have therefore espoused a strategy of participatory design [12], which early in the design process establishes a process of iterative refinement of solutions in partnership with users who, as the data show, are often impressively knowledgeable of complex sociotechnical concepts related to contemporary technologies. Our hope is that this methodological approach, besides reducing costly trial-and-error cycles, has a better chance of establishing long-term trust between researchers, administrators, and developers, leading to the sustained development and adoption of novel, useful, and usable informatics solutions for translational science.

Conclusions

We believe that our study makes four contributions. First, it provides direct and detailed evidence of the resource needs of CTS researchers. Second, it identifies concrete recommendations and cautions for the design of informatics solutions to help find and use computational and human resources. Third, it provides a semistructured interview instrument and an example of how it can be used to guide the design of contextually relevant informatics solutions. Fourth, it identifies potential deficiencies of well-known theoretical frameworks used to guide our study and suggestions for future remedies.

Our goal is to obtain feedback on our prototype using an iterative participatory design approach. Furthermore, we hope to conduct a survey to elicit the specific resource needs that should be included in the portal. The interview instrument and overall approach could therefore be used by other CTSA institutions to design current and future informatics solutions that are useful, usable, and contextually relevant to the populations that they serve, with the ultimate goal of accelerating progress in translational science.

Acknowledgments

We thank the participants in the study for their time and insights, and G Hiller, D Clauw, A Austin, K Smith, D Hunscher, A Atreja, G Omenn, and the anonymous reviewers for their valuable suggestions. This project was supported in part by grants UL1RR024986 and UL1RR029876 received from the National Center for Research Resources, a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured interview instrument.

[[PDF File \(Adobe PDF File\), 54KB](#) - [jmir_v14i3e75_app1.pdf](#)]

References

1. Zerhouni EA. Translational and clinical science--time for a new vision. *N Engl J Med* 2005 Oct 13;353(15):1621-1623 [[FREE Full text](#)] [doi: [10.1056/NEJMs053723](https://doi.org/10.1056/NEJMs053723)] [Medline: [16221788](#)]

2. National Center for Research Resources. 2011. NCRR Fact Sheet: Clinical and Translational Science Awards URL: https://www.ctsacentral.org/documents/Communication_Toolkit/CTSA_FactSheet.pdf [accessed 2012-01-09] [WebCite Cache ID 66J46gzx7]
3. McConnell C. National Center for Research Resources. 2009 Nov 2. NIH Awards \$27 Million in Recovery Act Funds to Enhance Scientist and Resource Networking URL: <http://www.nih.gov/news/health/nov2009/ncrr-02.htm> [accessed 2012-01-09] [WebCite Cache ID 66J49bWws]
4. Stanford University School of Medicine. 2012. CAP Network is a Virtual Workspace for Stanford University URL: <http://cap.stanford.edu/about.html> [accessed 2012-01-09] [WebCite Cache ID 66J4C2Xso]
5. Weber GM, Barnett W, Conlon M, Eichmann D, Kibbe W, Falk-Krzesinski H, Direct2Experts Collaboration. Direct2Experts: a pilot national network to demonstrate interoperability among research-networking platforms. *J Am Med Inform Assoc* 2011 Dec;18 Suppl 1:i157-i160. [doi: [10.1136/amiajnl-2011-000200](https://doi.org/10.1136/amiajnl-2011-000200)] [Medline: [22037890](https://pubmed.ncbi.nlm.nih.gov/22037890/)]
6. Schleyer T, Spallek H, Butler BS, Subramanian S, Weiss D, Poythress ML, et al. Facebook for scientists: requirements and services for optimizing how scientific collaborations are established. *J Med Internet Res* 2008;10(3):e24 [FREE Full text] [doi: [10.2196/jmir.1047](https://doi.org/10.2196/jmir.1047)] [Medline: [18701421](https://pubmed.ncbi.nlm.nih.gov/18701421/)]
7. Spallek H, Schleyer T, Butler BS. Good partners are hard to find: the search for and selection of collaborators in the health sciences. In: *Proceedings of eScience. 2008 Presented at: IEEE Fourth International Conference on eScience; Dec 7-12, 2008; Indianapolis, IN, USA.* [doi: [10.1109/eScience.2008.105](https://doi.org/10.1109/eScience.2008.105)]
8. Butler BS, Schleyer T, Spallek H. University of Pittsburgh. 2008 Mar 14. Evaluation and Redesign of an Index of Faculty Research Interests: Technical Report #031408 URL: http://www.dental.pitt.edu/informatics/orc/FRIP_report_031408.doc [accessed 2012-01-18] [WebCite Cache ID 66J4KYXDo]
9. Kaplan B, Harris-Salamone KD. Health IT success and failure: recommendations from literature and an AMIA workshop. *J Am Med Inform Assoc* 2009;16(3):291-299 [FREE Full text] [doi: [10.1197/jamia.M2997](https://doi.org/10.1197/jamia.M2997)] [Medline: [19261935](https://pubmed.ncbi.nlm.nih.gov/19261935/)]
10. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Manag Inf Syst Q* 1989;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
11. Barki H, Hartwick J. Measuring user participation, use involvement, and user attitude. *Manag Inf Syst Q* 1994;18:59-82. [doi: [10.2307/249610](https://doi.org/10.2307/249610)]
12. Schuler D, Namioka A. *Participatory Design: Principles and Practices*. Hillsdale, NJ: L. Erlbaum Associates; 1993.
13. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *Manag Inf Syst Q* 2003;27(3):425-478. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
14. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010 Feb;43(1):159-172. [doi: [10.1016/j.jbi.2009.07.002](https://doi.org/10.1016/j.jbi.2009.07.002)] [Medline: [19615467](https://pubmed.ncbi.nlm.nih.gov/19615467/)]
15. Grudin J. Computer-supported cooperative work: history and focus. *Computer* 1994;27(5):19-26. [doi: [10.1109/2.291294](https://doi.org/10.1109/2.291294)]
16. Olson GM, Olson JS. Distance matters. *Hum Comput Interact* 2000;15(2):139-178. [doi: [10.1207/S15327051HCI1523_4](https://doi.org/10.1207/S15327051HCI1523_4)]
17. Pratt W, Reddy MC, McDonald DW, Tarczy-Hornoch P, Gennari JH. Incorporating ideas from computer-supported cooperative work. *J Biomed Inform* 2004 Apr;37(2):128-137. [doi: [10.1016/j.jbi.2004.04.001](https://doi.org/10.1016/j.jbi.2004.04.001)] [Medline: [15120659](https://pubmed.ncbi.nlm.nih.gov/15120659/)]
18. Patton M. *Qualitative Evaluation and Research Methods*. Newbury Park, CA: Sage Publications; 1990.
19. Kvale S. *InterViews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, CA: Sage Publications; 1996.
20. Goodman LA. Snowball sampling. *Ann Math Stat* 1961;32(1):148-170.
21. Glaser BG, Strauss AL. *The Discovery of Grounded Theory*. New York: de Gruyter; 1967.
22. Tang PC, Patel VL. Major issues in user interface design for health professional workstations: summary and recommendations. *Int J Biomed Comput* 1994 Jan;34(1-4):139-148. [Medline: [8125627](https://pubmed.ncbi.nlm.nih.gov/8125627/)]
23. Pressman RS. *Software Engineering: A Practitioner's Approach*. New York, NY: McGraw-Hill; 1992.
24. Boyd D, Ellison N. Social network sites: definition, history, and scholarship. *J Comput Mediat Commun* 2007;13(1):210-230. [doi: [10.1111/j.1083-6101.2007.00393.x](https://doi.org/10.1111/j.1083-6101.2007.00393.x)]
25. The Java Community Process Program. 2011. JSR-000168 Portlet Specification (Final Release) URL: <http://jcp.org/aboutJava/communityprocess/final/jsr168/index.html> [accessed 2012-03-20] [WebCite Cache ID 66J5Fjqj2]
26. Oliver RL. A cognitive model of the antecedents and consequences of satisfaction decisions. *J Marketing Res* 1980;17(4):460-469. [doi: [10.2307/3150499](https://doi.org/10.2307/3150499)]
27. Cadotte ER, Woodruff RB, Jenkins RL. Expectations and norms in models of consumer satisfaction. *J Marketing Res* 1987;24(3):305-314. [doi: [10.2307/3151641](https://doi.org/10.2307/3151641)]
28. Halstead D. The use of comparison standards in customer satisfaction research and management: a review and proposed typology. *J Mark Theory Pract* 1999;7(3):13-26. [doi: [10.2307/40469965](https://doi.org/10.2307/40469965)]
29. Bhattacharjee A. Understanding information systems continuance: an expectation-confirmation model. *Manag Inf Syst Q* 2001;25(3):351-370. [doi: [10.2307/3250921](https://doi.org/10.2307/3250921)]

Abbreviations

CTSA: Clinical and Translational Science Awards
CTS: clinical and translational science

TAM: technology acceptance model

Edited by G Eysenbach; submitted 06.08.11; peer-reviewed by B Butler, S Lim Choi Keung; comments to author 07.09.11; revised version received 09.01.12; accepted 25.01.12; published 05.06.12.

Please cite as:

Bhavnani SK, Warden M, Zheng K, Hill M, Athey BD

Researchers' Needs for Resource Discovery and Collaboration Tools: A Qualitative Investigation of Translational Scientists

J Med Internet Res 2012;14(3):e75

URL: <http://www.jmir.org/2012/3/e75/>

doi: [10.2196/jmir.1905](https://doi.org/10.2196/jmir.1905)

PMID: [22668750](https://pubmed.ncbi.nlm.nih.gov/22668750/)

©Suresh K Bhavnani, Michael Warden, Kai Zheng, Mary Hill, Brian D Athey. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 05.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Mining Online Social Network Data for Biomedical Research: A Comparison of Clinicians' and Patients' Perceptions About Amyotrophic Lateral Sclerosis Treatments

Carlos Nakamura¹, PhD; Mark Bromberg², MD, PhD; Shivani Bhargava³, BA; Paul Wicks³, PhD; Qing Zeng-Treitler¹, PhD

¹Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

²Department of Neurology, University of Utah, Salt Lake City, UT, United States

³PatientsLikeMe, Cambridge, MA, United States

Corresponding Author:

Carlos Nakamura, PhD

Department of Biomedical Informatics

University of Utah

Room 5775 HSEB

26 South 2000 East

Salt Lake City, UT, 84112-5775

United States

Phone: 1 801 213 3357

Fax: 1 801 581 4297

Email: carlos.nakamura@utah.edu

Abstract

Background: While only one drug is known to slow the progress of amyotrophic lateral sclerosis (ALS), numerous drugs can be used to treat its symptoms. However, very few randomized controlled trials have assessed the efficacy, safety, and side effects of these drugs. Due to this lack of randomized controlled trials, consensus among clinicians on how to treat the wide range of ALS symptoms and the efficacy of these treatments is low. Given the lack of clinical trials data, the wide range of reported symptoms, and the low consensus among clinicians on how to treat those symptoms, data on the prevalence and efficacy of treatments from a patient's perspective could help advance the understanding of the symptomatic treatment of ALS.

Objective: To compare clinicians' and patients' perspectives on the symptomatic treatment of ALS by comparing data from a traditional survey study of clinicians with data from a patient social network.

Methods: We used a survey of clinicians' perceptions by Forsheew and Bromberg as our primary data source and adjusted the data from PatientsLikeMe to allow for comparisons. We first extracted the 14 symptoms and associated top four treatments listed by Forsheew and Bromberg. We then searched the PatientsLikeMe database for the same symptom-treatment pairs. The PatientsLikeMe data are structured and thus no preprocessing of the data was required.

Results: After we eliminated pairs with a small sample, 15 symptom-treatment pairs remained. All treatments identified as useful were prescription drugs. We found similarities and discrepancies between clinicians' and patients' perceptions of treatment prevalence and efficacy. In 7 of the 15 pairs, the differences between the two groups were above 10%. In 3 pairs the differences were above 20%. Lorazepam to treat anxiety and quinine to treat muscle cramps were among the symptom-treatment pairs with high concordance between clinicians' and patients' perceptions. Conversely, amitriptyline to treat labile emotional effect and oxybutynin to treat urinary urgency displayed low agreement between clinicians and patients.

Conclusions: Assessing and comparing the efficacy of the symptomatic treatment of a complex and rare disease such as ALS is not easy and needs to take both clinicians' and patients' perspectives into consideration. Drawing a reliable profile of treatment efficacy requires taking into consideration many interacting aspects (eg, disease stage and severity of symptoms) that were not covered in the present study. Nevertheless, pilot studies such as this one can pave the way for more robust studies by helping researchers anticipate and compensate for limitations in their data sources and study design.

(*J Med Internet Res* 2012;14(3):e90) doi:[10.2196/jmir.2127](https://doi.org/10.2196/jmir.2127)

KEYWORDS

Amyotrophic lateral sclerosis; ALS; Lou Gehrig's disease; treatment response; treatment efficacy; online social network; online patient community

Introduction

Clinical trials, surveys, and medical records are conventional sources of data for medical research. Some researchers have begun to explore alternative data sources to better understand the patient's perspective [1]. An example of such alternative data sources is online social networks such as PatientsLikeMe [2]. In recent years, online social network sites with a health focus have attracted large numbers of users and have begun accumulating large quantities of detailed clinical information. Information gathered by these networks is primarily intended for patients to share with each other. However, such information has also begun to attract the attention of medical researchers for a variety of uses including clinical trial recruitment [3], development of patient-reported outcomes [4,5], and perhaps even novel ways of evaluating treatment efficacy [6-8].

Because social networks are a relatively new phenomenon, their strengths and limitations as a data source for medical research have not been carefully investigated. On the one hand, the large amount of patient data on social networks could bring a great contribution to studies on important issues such as treatment response and compliance. On the other hand, we need to understand whether and how these data provide us with additional information in comparison with traditional data sources, as well as the validity of that information.

In this study, we proposed to compare clinicians' and patients' perspectives on the symptomatic treatment of amyotrophic lateral sclerosis (ALS) by comparing data from a traditional survey study of clinicians with data from a patient social network. Forsheew and Bromberg extracted data on clinicians' perspectives from a national survey among ALS clinical centers [9]. Data on patients' perspectives were extracted from the online social network PatientsLikeMe.

ALS, also known as Lou Gehrig's disease, is a rare but fatal neurodegenerative disease that causes muscle weakness. Most patients die of respiratory failure within 3 to 5 years of the disease onset [10].

While only one drug is known to slow the progress of ALS (riluzole), numerous drugs can be used to treat its diverse array of symptoms. However, very few randomized controlled trials have assessed the efficacy, safety, and side effects of potential treatments to modulate these symptoms. Consensus among clinicians on how to treat the wide range of ALS symptoms and the efficacy of these treatments have historically been low [9]. The American Academy of Neurology, in developing their recent guidelines on the management of ALS [11], found only limited evidence to support treatment choice for symptom management. Evidence was most robust for treatment of excessive saliva (level B evidence for botox, level C evidence for irradiation of the salivary gland) and pseudobulbar affect (level B evidence for the efficacy of dextromethorphan and quinidine), but wholly lacking for treatment of fatigue (level C evidence for withdrawing riluzole), cramps (level U, no

appropriate studies), spasticity (level U), depression (level U), and anxiety (level U). Given (1) the lack of evidence to support decisions (2), the wide range of reported symptoms, and (3) the low consensus among clinicians on how to treat those symptoms, data on the prevalence and efficacy of treatments from a patient's perspective could help advance the understanding of the symptomatic treatment of ALS.

Background

One approach to assessing the efficacy of symptomatic drugs when randomized controlled trials are lacking is through a survey instrument fielded to clinicians, patients, or caregivers. In 2003, Forsheew and Bromberg [9] reported the results of a national survey sent out to 61 ALS clinics in North America, to which 39 clinics replied. Clinicians were asked to name the three most frequent interventions they would use for each symptom and their respective rankings for their perception of efficacy in relieving the symptom (ranging from 1, rarely effective, to 5, nearly complete relief). The 14 symptoms included in the survey were selected based on a literature search. The article reported the four top interventions for each of the predetermined symptoms.

Forsheew and Bromberg mentioned in their article that a mirror study giving the patients' perspective on the same topic was underway, though the results of that study have not been published yet. In the absence of such a study, it is still possible to contrast clinicians' and patients' perspectives by looking into rapidly increasing and publicly available databases of online communities of patients. PatientsLikeMe is an online community that allows members to track their progress with clinical scales, share information, and learn more about their condition. It has established data-sharing partnerships with doctors, pharmaceutical and medical device companies, research organizations, and nonprofits. Starting with an ALS patient community, PatientsLikeMe now hosts communities for more than 100,000 patients with any medical condition. Currently, the US ALS user base represents about 5% of the US ALS population and is an almost perfect match in terms of demographics, but with a bias toward long-term survivors. By definition, a bias toward patients with computer and Web access is indeed also present.

Each community member is asked to track 11 to 12 primary symptoms of his or her condition. The primary symptom list was generated for the ALS community with input from health care practitioners and the literature. Members can also report, in free-text form, any additional symptoms they are experiencing. The result is a semistructured list, which patients can use as an aid for future symptom reporting. It also permits comparison with symptoms reported by other patients in the system. Additional symptoms entered by a member lead to a prompt ("did you mean...") to attempt to merge their description with existing symptom entries; if they wish to add a new symptom term, this is reviewed by a nurse curator.

Comparing data from very different sources is challenging, as the many adjustments that datasets must go through may induce further biases. However, in the absence of robust evidence from randomized controlled trials (which also have their own limitations and weaknesses), the data from surveys and online communities of patients become a viable resource in the search for a better understanding of clinicians' and patients' perceptions.

Methods

Treatment Prevalence

We used the Forsheew and Bromberg study as our primary data source and adjusted the data from PatientsLikeMe to allow for comparisons. We first extracted the 14 symptoms and associated top four treatments listed by Forsheew and Bromberg. We then searched the PatientsLikeMe database for the same symptom–treatment pairs. Members of PatientsLikeMe also mentioned 13 of the 14 symptoms listed by Forsheew and Bromberg. The only nonoverlapping symptom was thick phlegm, which had no corresponding item. The PatientsLikeMe data are structured and thus no preprocessing of the data was required.

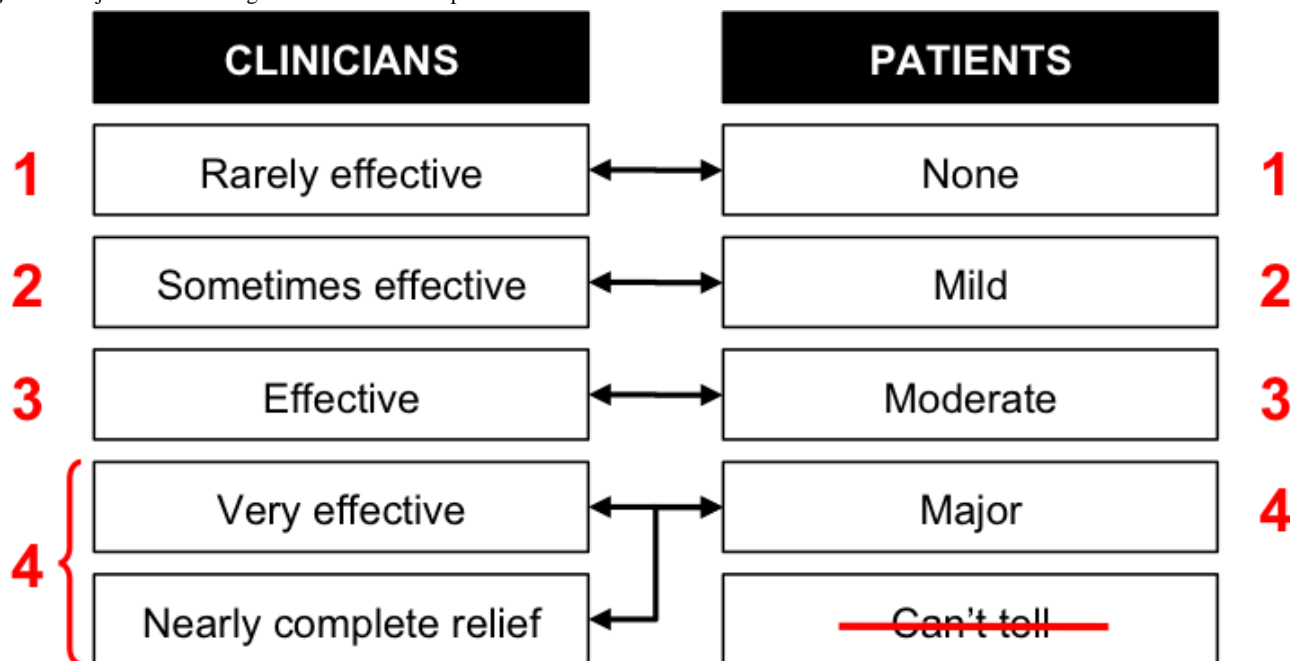
Treatment Efficacy

In this study we proposed to compare clinicians' and patients' perceptions of the efficacy of different treatments to alleviate ALS symptoms. There is one important distinction between the perceptions of these two groups. Patients' perception of treatment efficacy is direct and personal. That is, patients evaluate treatment efficacy based on how well they responded

to the treatment. Clinicians' perception, on the other hand, is an aggregate judgment, but presumably is based on how well the treatment worked for their patient population.

To enable a fair comparison between the two datasets, we made a few adjustments and transformations. PatientsLikeMe provided a dataset for analysis. We started by searching for efficacy data that would match the 56 symptom–treatment pairs from the Forsheew and Bromberg study. We then discarded data on any symptom–treatment pair for which patient sample size was smaller than 6 patients. Given that each pair had a different sample size, we converted the frequencies to percentages to facilitate a visual comparison. Finally, we adjusted the rating scales from both sources. The rating scales from the Forsheew and Bromberg study and from PatientsLikeMe were composed of five levels. Items rated as “Can't tell” from the PatientsLikeMe dataset were not included in our comparative analysis because (1) they were not part of an ordinal scale like the 4 other items (ie, none, mild, moderate, or major), and (2) they referred to cases in which patients had not yet formed an opinion. We combined the two top levels in the Forsheew and Bromberg study because the last item in the scale, nearly complete relief, was quite small (corresponding to less than 3% of the ratings), and it interrupted the continuity of the scale. The first 4 items in the scale explicitly used the word *effective*. Further, the first two qualifiers (ie, *rarely* and *sometimes*) strongly suggested a rating scale based on frequency. The last item in the scale, however, suggested a rating based on the intensity rather than on the frequency of the effect by using the qualifier *complete*. Figure 1 summarizes the adjustments made to both scales to facilitate a comparative analysis.

Figure 1. Adjustment of rating scales to enable comparison.



Results

As of March 14, 2010, the ALS community on PatientsLikeMe consisted of 4375 members, of whom 41% were female and 59% were male. Of the 697 members who disclosed their age,

3% were under 30 years old, 10% were between 30 and 39 years old, 24% were between 40 and 49 years old, 29% were between 50 and 59 years old, 26% were between 60 and 69 years old, and 8% were older than 69 years. This gender and age distribution is consistent with that of the general ALS

population: ALS occurs more frequently in men than in women with a ratio of approximately 1.5:1 to 2.0:1. It is also a disease of upper middle age with a mean age of onset between 55 and 65 years [12].

Prevalence of Treatment Usage

After filtering out the symptoms with fewer than 10 respondents from the list of 13 overlapping symptoms, 10 symptoms and 15 symptom–treatment pairs remained. Table 1 shows the

number of patients who responded to the question on the treatments prescribed or received for the 10 specific symptoms. Table 2 shows the 15 symptom–treatment pairs. None of the 15 pairs was used by either group more than 50% of the time. In 7 of the 15 pairs, the discrepancy in usage between the two groups was above 10%. In 3 pairs the differences were over 20%: amitriptyline for labile emotional effect; oxybutynin for urinary urgency; and amitriptyline for disturbed sleep.

Table 1. Number of clinicians and patients who responded to the question regarding treatments prescribed or received for specific symptoms.

Symptom	Clinicians' report	Patients' report
Muscle cramps	31 (51%)	95 (35%)
Sialorrhea	30 (49%)	258 (17%)
Depression	31 (51%)	300 (19%)
Anxiety	29 (48%)	240 (15%)
Spasticity and stiffness	31 (51%)	848 (39%)
Labile emotional effect	28 (46%)	147 (9%)
Urinary urgency	23 (38%)	31 (25%)
Disturbed sleep	31 (51%)	216 (15%)
Fasciculations	20 (33%)	112 (5%)
Constipation	28 (46%)	200 (15%)

Table 2. Comparison of percentage of clinicians reporting preferences for specific treatment use in response to a symptom, and patient-reported usage of the same treatments.

Symptom	Treatment	Clinicians' report	Patients' report
Muscle cramps	Quinine	30 (35%)	30 (32%)
	Baclofen	16 (19%)	26 (27%)
Sialorrhea	Amitriptyline	32 (36%)	75 (29%)
	Glycopyrrolate	20 (23%)	46 (18%)
Depression	Sertraline	21 (24%)	50 (17%)
	Fluoxetine	16 (18%)	28 (9%)
Anxiety	Lorazepam	16 (24%)	65 (27%)
Spasticity and stiffness	Baclofen	37 (40%)	246 (29%)
	Tizanidine	21 (23%)	42 (5%)
Labile emotional effect	Amitriptyline	30 (50%)	31 (21%)
Urinary urgency	Oxybutynin	22 (46%)	6 (19%)
Disturbed sleep	Amitriptyline	29 (34%)	24 (11%)
	Zolpidem	14 (16%)	71 (33%)
Fasciculations	Gabapentin	10 (20%)	9 (8%)
Constipation	Docusate	13 (16%)	38 (19%)

Of the symptom–treatment pairs, 13 corresponded to US Food and Drug Administration-approved (*indicated*) uses of the drugs; 2 pairs corresponded to uses for which an indication had not been granted (*off-label use*): amitriptyline for labile emotional effect and gabapentin for fasciculations.

Treatment Efficacy

After eliminating symptom–treatment pairs for which the number of efficacy evaluations was smaller than 10, we were left with 5 pairs. In the following paragraphs we present the results for the perceived efficacy of amitriptyline to treat sialorrhea and baclofen to treat spasticity and stiffness. Results

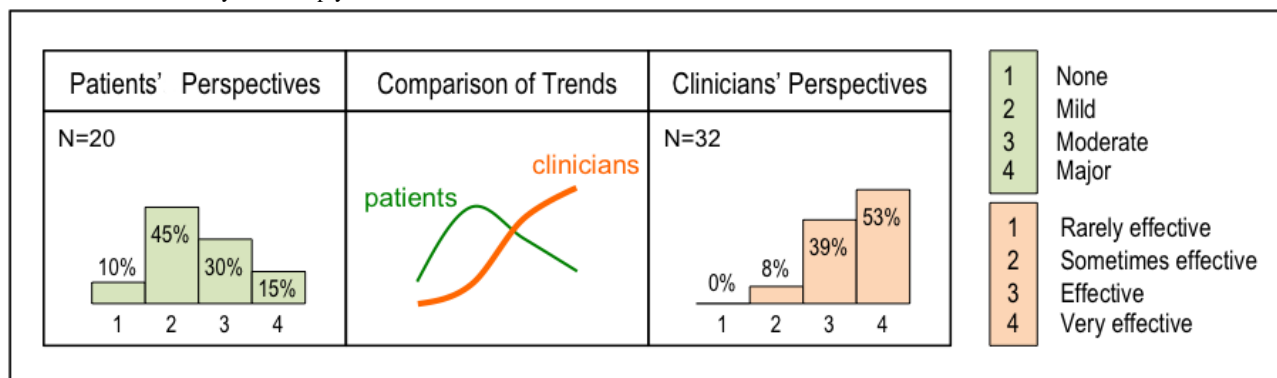
for the remaining 3 pairs are provided in [Multimedia Appendix 1](#).

Perceived Efficacy of Amitriptyline to Treat Sialorrhea

Clinicians had a more positive perception of the efficacy of amitriptyline ([Figure 2](#)) to treat sialorrhea. Clinician's ratings

for the drug show a clear ascending trend from ratings 2 to 4, with no clinicians' ratings of 1. Patients' ratings of amitriptyline show an inverse trend, with a positively skewed distribution.

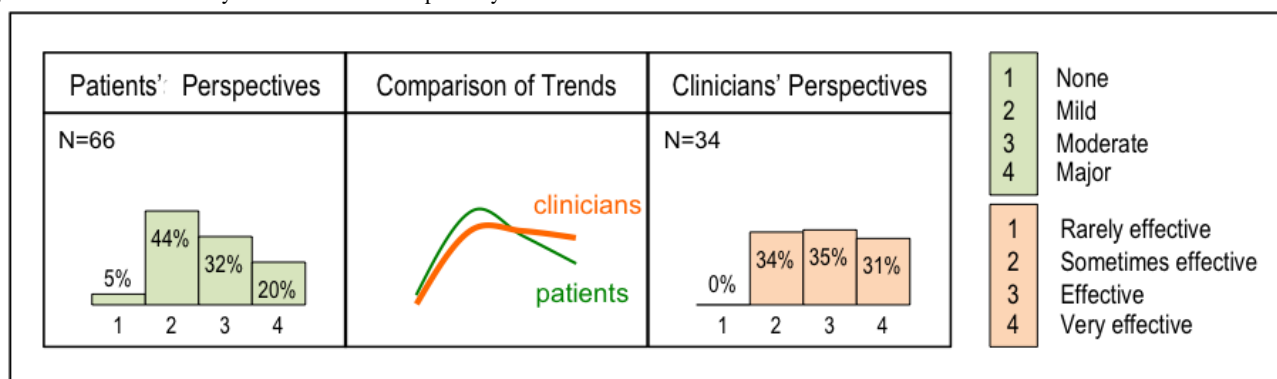
Figure 2. Perceived efficacy of amitriptyline to treat sialorrhea.



Perceived Efficacy of Baclofen to Treat Spasticity and Stiffness

Patients' ratings for baclofen showed a positively skewed frequency distribution ([Figure 3](#)). Clinicians' ratings were evenly distributed between categories 2 and 4.

Figure 3. Perceived efficacy of baclofen to treat spasticity and stiffness.



Discussion

We found that clinicians' and patients' perceptions of treatment prevalence and efficacy had some differences and could potentially complement each other. Both patients and clinicians nominated a wide range of treatments for ALS symptoms, and there was not always consensus between them regarding the efficacy of each individual treatment. This indicates that more robust empirical data to guide clinical practices in the management of ALS symptoms is required, and the data from the social network may be used to inform practice.

The clinicians perceived treatment efficacy as only marginal for some treatments. Symptoms associated with treatments that received an average ranking of 3 or higher were classified as treatable. Symptoms associated with treatments that scored below that threshold were classified as difficult to treat. Patients sometimes give a more positive ranking and sometimes a more

negative ranking. Due to the sample size, we did not statistically test the differences.

Limited studies have been done to assess the value of using social network data for research purposes. We are not aware of prior work that compared data gathered from a social network with data from clinicians. Observing consistencies and inconsistencies between the two data sources in this study, we believe that data gathered from an online social network, especially as it grows in popularity, can be a potential data source of patient perspectives for scientific studies.

Limitations and Future Studies

While providing valuable data, Forshaw and Bromberg also duly acknowledge one limitation in the design of their survey: respondents ranked treatments based on an average ALS patient, which oversimplifies the real issue, given that response to treatment may vary according to the disease stage.

Both clinician and patient cohorts are expected to be heavily skewed by self-selection bias. That is, the sample is less likely to be a good representation of a larger population when it is not completely random. In fact, complete sample randomization is a constant issue in any study that deals with human participants. We do not know the demographics or clinical characteristics of patients treated by the physician participants in the Forsheew and Bromberg study. We can assume that the patient population was roughly representative of the general patient population, since Forsheew and Bromberg conducted a national survey. Similarly, we have little knowledge of the PatientsLikeMe population's clinical characteristics. The size of the ALS community on PatientsLikeMe, however, is fairly large, representing roughly 5% of the US ALS population. The age and gender of the PatientsLikeMe ALS population do resemble those of the general ALS population.

Given the very small sample sizes for both datasets, the outcomes regarding treatment efficacy are not as robust as one would wish. However, given that the patient pool of PatientsLikeMe grows daily, we expect that such inaccuracies due to sample size will become increasingly smaller in the case of patient-reported data.

In terms of relative sample size, cohorts of clinicians and patients do not lend themselves to a direct comparison. Patients' observations are based on their personal experiences and thus are direct and singular. Clinicians' observations, on the other hand, are indirect and aggregate—that is, they are formed based on data coming from a collection of patients. Further, clinicians obviously possess more extensive clinical knowledge than patients do, which makes clinicians' observations more systematic. Consequently, studies based on patients' perspectives are likely to require larger sample sizes than studies based on clinicians' perspectives.

In this study we compared data from only one survey and one online community of patients. Future studies should consider including more instances of each type of data source, which would compensate for the specific limitations and weaknesses of each data source.

Several authors have called attention to the fact that self-selection bias is a common threat to the internal and external validity in studies involving human participants [12,13]. This threat is further magnified in studies that use measurements based on participants' perception.

Conclusions

Our comparative analysis detected significant discrepancies between patients' and clinicians' perceptions of treatment prevalence and efficacy. Both data sets have their own limitations, which were partially described in the previous sections. However, if they are properly conjugated, their potential to offer a more complete and accurate picture of ALS symptomatic treatment is likely to increase. As previously mentioned, data validity issues related to sample size in PatientsLikeMe are likely to gradually self-correct as the online community grows.

Assessing and comparing the efficacy of the symptomatic treatment of a complex and rare disease such as ALS is not easy. It is also important that patients' as well as clinicians' views be taken into consideration. Drawing a reliable profile of treatment efficacy requires taking into consideration many interacting aspects (eg, disease stage and severity of symptoms) that were not covered in the present study. Nevertheless, pilot studies such as the one described here can pave the way for more robust studies by helping researchers anticipate and compensate for limitations in their data sources and study design.

Acknowledgments

This work was supported by National Institutes of Health grant R21 NS067463.

Conflicts of Interest

PW and SB are employees of PatientsLikeMe and hold stock options in the company. The PatientsLikeMe research and development team has received research funding from Abbott, Accorda, Avanir, Biogen, Genzyme, Merck, Novartis, Sanofi, and UCB.

Multimedia Appendix 1

Perceived efficacy of other treatments.

[PDF File (Adobe PDF File), 28KB - [jmir_v14i3e90_app1.pdf](#)]

References

1. ALSUntangled Group. ALSUntangled No. 8: Low dose naltrexone for ALS. *Amyotroph Lateral Scler* 2011 Jan;12(1):76-78. [doi: [10.3109/17482968.2010.544386](#)] [Medline: [21174518](#)]
2. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. *J Med Internet Res* 2010;12(2):e19 [FREE Full text] [doi: [10.2196/jmir.1549](#)] [Medline: [20542858](#)]
3. Allison M. Can web 2.0 reboot clinical trials? *Nat Biotechnol* 2009 Oct;27(10):895-902. [doi: [10.1038/nbt1009-895](#)] [Medline: [19816442](#)]

4. Wicks P, Massagli M, Kulkarni A, Dastani H. Use of an online community to develop patient-reported outcome instruments: the Multiple Sclerosis Treatment Adherence Questionnaire (MS-TAQ). *J Med Internet Res* 2011;13(1):e12 [FREE Full text] [doi: [10.2196/jmir.1687](https://doi.org/10.2196/jmir.1687)] [Medline: [21266318](https://pubmed.ncbi.nlm.nih.gov/21266318/)]
5. Wicks P, Massagli MP, Wolf C, Heywood J. Measuring function in advanced ALS: validation of ALSFRS-EX extension items. *Eur J Neurol* 2009 Mar;16(3):353-359. [doi: [10.1111/j.1468-1331.2008.02434.x](https://doi.org/10.1111/j.1468-1331.2008.02434.x)] [Medline: [19364363](https://pubmed.ncbi.nlm.nih.gov/19364363/)]
6. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat Biotechnol* 2011 May;29(5):411-414. [doi: [10.1038/nbt.1837](https://doi.org/10.1038/nbt.1837)] [Medline: [21516084](https://pubmed.ncbi.nlm.nih.gov/21516084/)]
7. Frost J, Okun S, Vaughan T, Heywood J, Wicks P. Patient-reported outcomes as a source of evidence in off-label prescribing: analysis of data from PatientsLikeMe. *J Med Internet Res* 2011;13(1):e6 [FREE Full text] [doi: [10.2196/jmir.1643](https://doi.org/10.2196/jmir.1643)] [Medline: [21252034](https://pubmed.ncbi.nlm.nih.gov/21252034/)]
8. Moreno MA, Parks MR, Zimmerman FJ, Brito TE, Christakis DA. Display of health risk behaviors on MySpace by adolescents: prevalence and associations. *Arch Pediatr Adolesc Med* 2009 Jan;163(1):27-34. [doi: [10.1001/archpediatrics.2008.528](https://doi.org/10.1001/archpediatrics.2008.528)] [Medline: [19124700](https://pubmed.ncbi.nlm.nih.gov/19124700/)]
9. Forshew DA, Bromberg MB. A survey of clinicians' practice in the symptomatic treatment of ALS. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2003 Dec;4(4):258-263. [Medline: [14753660](https://pubmed.ncbi.nlm.nih.gov/14753660/)]
10. Wijesekera LC, Leigh PN. Amyotrophic lateral sclerosis. *Orphanet J Rare Dis* 2009;4:3 [FREE Full text] [doi: [10.1186/1750-1172-4-3](https://doi.org/10.1186/1750-1172-4-3)] [Medline: [19192301](https://pubmed.ncbi.nlm.nih.gov/19192301/)]
11. Miller RG, Jackson CE, Kasarskis EJ, England JD, Forshew D, Johnston W, Quality Standards Subcommittee of the American Academy of Neurology. Practice parameter update: the care of the patient with amyotrophic lateral sclerosis: multidisciplinary care, symptom management, and cognitive/behavioral impairment (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 2009 Oct 13;73(15):1227-1233. [doi: [10.1212/WNL.0b013e3181bc01a4](https://doi.org/10.1212/WNL.0b013e3181bc01a4)] [Medline: [19822873](https://pubmed.ncbi.nlm.nih.gov/19822873/)]
12. Dorland WA. *Dorland's Illustrated Medical Dictionary*. Philadelphia, PA: Saunders; 1988.
13. Lieu JE, Dewan K. Assessment of self-selection bias in a pediatric unilateral hearing loss study. *Otolaryngol Head Neck Surg* 2010 Mar;142(3):427-433. [doi: [10.1016/j.otohns.2009.11.035](https://doi.org/10.1016/j.otohns.2009.11.035)] [Medline: [20172393](https://pubmed.ncbi.nlm.nih.gov/20172393/)]

Abbreviations

ALS: amyotrophic lateral sclerosis

Edited by G Eysenbach; submitted 03.04.12; peer-reviewed by J Luo; comments to author 25.04.12; revised version received 23.05.12; accepted 01.06.12; published 21.06.12.

Please cite as:

Nakamura C, Bromberg M, Bhargava S, Wicks P, Zeng-Treitler Q

Mining Online Social Network Data for Biomedical Research: A Comparison of Clinicians' and Patients' Perceptions About Amyotrophic Lateral Sclerosis Treatments

J Med Internet Res 2012;14(3):e90

URL: <http://www.jmir.org/2012/3/e90/>

doi: [10.2196/jmir.2127](https://doi.org/10.2196/jmir.2127)

PMID: [22721865](https://pubmed.ncbi.nlm.nih.gov/22721865/)

©Carlos Nakamura, Mark Bromberg, Shivani Bhargava, Paul Wicks, Qing Zeng-Treitler. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 21.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Use of Social Media by Western European Hospitals: Longitudinal Study

Tom H Van de Belt¹, MSc; Sivera AA Berben², RN, PhD; Melvin Samsom³, MD, PhD; Lucien JLPG Engelen^{1,2}; Lisette Schoonhoven^{2,4,5}, RN, PhD

¹Radboud REshape and Innovation Centre, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

²Emergency Healthcare Network, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

³Executive Board, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

⁴Scientific Institute for Quality of Healthcare, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

⁵Faculty of Health Sciences, University of Southampton, Southampton, United Kingdom

Corresponding Author:

Tom H Van de Belt, MSc

Radboud REshape and Innovation Centre

Radboud University Nijmegen Medical Centre

1st Floor

St Annastraat 299

Nijmegen, 6512 GT

Netherlands

Phone: 31 24 7440083

Fax: 31 24 7440083

Email: t.vandebelt@reshape.umcn.nl

Abstract

Background: Patients increasingly use social media to communicate. Their stories could support quality improvements in participatory health care and could support patient-centered care. Active use of social media by health care institutions could also speed up communication and information provision to patients and their families, thus increasing quality even more. Hospitals seem to be becoming aware of the benefits social media could offer. Data from the United States show that hospitals increasingly use social media, but it is unknown whether and how Western European hospitals use social media.

Objective: To identify to what extent Western European hospitals use social media.

Methods: In this longitudinal study, we explored the use of social media by hospitals in 12 Western European countries through an Internet search. We collected data for each country during the following three time periods: April to August 2009, August to December 2010, and April to July 2011.

Results: We included 873 hospitals from 12 Western European countries, of which 732 were general hospitals and 141 were university hospitals. The number of included hospitals per country ranged from 6 in Luxembourg to 347 in Germany. We found hospitals using social media in all countries. The use of social media increased significantly over time, especially for YouTube (n = 19, 2% to n = 172, 19.7%), LinkedIn (n = 179, 20.5% to n = 278, 31.8%), and Facebook (n = 85, 10% to n = 585, 67.0%). Differences in social media usage between the included countries were significant.

Conclusions: Social media awareness in Western European hospitals is growing, as well as its use. Social media usage differs significantly between countries. Except for the Netherlands and the United Kingdom, the group of hospitals that is using social media remains small. Usage of LinkedIn for recruitment shows the awareness of the potential of social media. Future research is needed to investigate how social media lead to improved health care.

(*J Med Internet Res* 2012;14(3):e61) doi:[10.2196/jmir.1992](https://doi.org/10.2196/jmir.1992)

KEYWORDS

Social media; health 2.0; medicine 2.0; eHealth; participatory health care; patient empowerment; Web 2.0; patient-centered care

Introduction

Social media are defined as a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and they allow the creation and exchange of user-generated content [1]. Social media allow individuals to participate in online social networks and turn communication into interactive dialogue, using highly accessible and scalable communication applications [2]. Of all young Internet users (18–24 years of age) in the European Union, 80% use social media [3]. In the Netherlands, this percentage is even higher, with 91% using social media [3].

Facebook and Twitter are well-known examples of social media, which have become mainstream social technologies [4]. Facebook has over 800 million active users [5]. For comparison, the United States has 310 million inhabitants [6]. One of the success factors of social media is that many are free of charge. Social media play an increasingly important role in our society, and they are being used for a large variety of purposes, varying from finding a job or an employee to finding a partner or planning a trip. Also, a growing number of people are using mobile devices such as smartphones and tablet computers, which allow them to use social media from any place, at any time [1].

Social media empower users by allowing them to communicate effectively and have access to all kinds of information. Not only individuals use social media; companies use them too. It helps them to listen better to customers to hear what they want. Barnes and Mattson studied use of blogs and Twitter by the 500 largest corporations in the United States [7]. They found a steady adoption of blogs and an explosive growth of Twitter. As these companies have great influence on the commercial sector, it is expected that social media will become more important in the business world.

In health care, patients increasingly use social media to communicate and share information. This is one of the fundamentals of what is described as Health 2.0 or Medicine 2.0 [8]. Patients share their stories and information on social media, which are rapidly indexed by search engines like Google and can be found easily. Seeing that many patients start by performing a Google search, it seems relevant for hospital organizations to be active on social media. For example, 64% of all respondents of an online questionnaire among patients in the United States start by performing a search to analyze their condition [9]. Another reason why hospital organizations should embrace social media is that it may contribute to quality improvements in health care. Active use of social media not only speeds up communication and improves information provision for patients; it allows caregivers to engage patients in the delivery of care, and for caregivers and patients to make decisions collaboratively and improve their relationship [10]. In this way, using social media improves patient-centered care [11].

There are also beneficial aspects for the hospital organization itself. Several studies reported that social media can improve communication among staff, facilitate networking, attract visitors to the hospital's website, build the hospital's brand, and be used for recruitment for research projects [12,13].

A descriptive study performed in the United Kingdom found that 40% of the 152 health care organizations they studied used one or more types of social media, but that there was little interaction with online visitors (eg, patients) [14]. Also, many organizations were simply “seeding” information. In the United States, the use of social media by hospitals has been noted. Bennett documented that 674 hospitals had a Twitter account and 448 were on YouTube [15]. Considering that the United States has a total of 5000 hospitals, around 15.7% of all hospitals in the United States are on Twitter, 20.3% are on Facebook, and 10.9% are on YouTube [16]. However, it is unknown whether and how Western European hospitals use social media. Therefore, the target of this study was to identify the extent to which European hospitals use social media.

Methods

In this longitudinal study, we explored the use of social media by hospitals in 12 Western European countries through an Internet search.

Inclusion Criteria

We included the following Western European countries: the Netherlands, Belgium, Luxembourg, Germany, Austria, Switzerland, the United Kingdom, Ireland, Norway, Sweden, Finland, and Denmark. To retrieve a comprehensive list of hospitals for each country, we searched for lists of hospitals with detailed information on Wikipedia and the Hospitals Worldwide website [17,18]. Second, we contacted colleagues from the included countries and asked for official lists of hospitals. Third, we consulted country-specific websites with detailed information. Fourth and last, we used Google and each hospital's website to find additional information such as contact information or the number of beds. We included only hospitals with a website and at least 200 beds. If hospitals were part of a larger hospital organization with a central website, we explored the central website only and counted these hospitals as 1 hospital.

Variables

For each hospital we recorded the following characteristics: official name, address, country, province or state, email, number of beds, and number of hospitals included in the organization.

Since no scientific evidence was available on the popularity of different social media, we used information from websites and infographics to decide which social media were most popular and needed to be included in the study [19,20]. We gathered data about the following social media: YouTube, Twitter, Facebook, LinkedIn, and blogs (weblogs). We defined blog by the presence of the following characteristics: reverse chronological order of publication, regular updates (>1 per month), and the possibility to post comments. Facebook has different types of pages. In this study, we distinguished between company pages and group pages. For each medium, we searched for relevant data on use such as the number of friends or followers, the number of videos or tweets, and the date of registration. For each medium, we recorded whether the media could be found via the hospital's website.

Data Collection

Between April 2009 and July 2011, we collected data for each country during the following three time periods: T1 (April to August 2009), T2 (August to December 2010), and T3 (April to July 2011). YouTube accounts, Twitter accounts, and blogs were measured at T1, T2, and T3. For Facebook and LinkedIn, we performed two measurements, at T2 and T3.

Two researchers collected the data. A predefined search protocol was used, containing a 3-step search strategy. First, we visited

the hospital's website and searched for social media. We also used the website's search function (if available). Second, we searched for the hospital's name within the different types of social media such as YouTube. Third and last, we used Google for more specific search queries, such as the hospital's name and *Twitter*. Table 1 presents the search protocol. Before the official start, the two researchers involved in the search discussed the results for 20 hospitals. Since all variables in this study are unambiguous (eg, number of beds, Twitter account: "yes" or "no"), no relevant differences or issues appeared.

Table 1. Search protocol for data collection.

Step	Protocol
1	Select hospital from list.
2	Visit official website and add contact information to table. Find using standard search tool (ie, Google).
3	Record number of beds (total). Include hospitals with >200 beds.
4	If included, proceed to next steps.
5	Add general information.
6	Look for different types of social media on hospital's website and add to table.
7	Use search option on hospital's website and search the terms <i>YouTube</i> , <i>movie</i> , <i>film</i> , <i>Twitter</i> , <i>Facebook</i> , <i>blog</i> , <i>LinkedIn</i> , and <i>weblog</i> . Add all new social media to the table.
8	Visit Twitter.com, Youtube.com, Facebook.com, and LinkedIn.com and search on hospital's official name. Add all new social media to the table.
9	Use specific search queries in Google, eg, the hospital's name AND <i>Facebook</i> . Add all new social media to the table.
10	Add other relevant information for all types of social media, eg, number of friends and followers, date of registration.

Data Validation

We contacted all organizations with the request to validate the results for their hospital. We sent emails to each hospital's general email address as stated on their official website, most likely on the Contact page. The email contained a description of this study by the Radboud REshape & Innovation Center, University Nijmegen Medical Centre, and a unique link to an online database. Receivers were able to make changes or add information or comments. We sent 873 email requests. Of these, 45 messages (5%) were returned as undeliverable, and 44 hospitals validated the results (5%).

Analysis

We used descriptive statistics to describe the basic features of our data and the use of social media by the included hospitals. We calculated percentages, means, and standard deviations for normally distributed data, and medians and interquartile ranges

for nonnormally distributed data. Cochran Q test was used to analyze the differences in social media usage between the three measurements within individual countries. In case of significant differences, we used the McNemar test for post hoc testing. Furthermore, we analyzed the differences in social media usage between countries at T3 by using the chi-square test. Finally, we used the Wilcoxon rank test to analyze the nonnormally distributed data for number of videos, views, and followers between T2 and T3 within the included countries.

Results

In total we looked at 873 hospitals from 12 Western European countries: 732 general hospitals and 141 university hospitals. The number of included hospitals per country ranged from 6 in Luxembourg to 347 in Germany. The mean number of beds per hospital was 544. Table 2 presents general characteristics of the hospitals.

Table 2. Hospitals included in the analysis and their general details.

Country ^a	Number of hospitals			Number of beds, mean (SD)
	Total	General hospitals	University hospitals	
NL	88	80	8	549 (278)
BE	91	79	12	450 (261)
LU	6	5	1	363 (139)
DE	347	314	33	533 (445)
AT	25	19	6	775 (587)
CH	41	39	2	389 (232)
UK	175	123	52	624 (282)
IR	28	21	7	392 (192)
NO	17	11	6	480 (238)
SE	22	17	5	698 (511)
FI	9	7	2	697 (544)
DK	24	17	7	551 (286)
Total	873	732	141	544 (376)

^a NL = the Netherlands, BE = Belgium, LU = Luxembourg, DE = Germany, AT = Austria, CH = Switzerland, UK = United Kingdom, IR = Ireland, NO = Norway, SE = Sweden, FI = Finland, DK = Denmark.

We found hospitals using social media in all countries. The use of social media increased over time, and we found significant differences between countries. Table 3, Table 4, and Table 5 show the results by country.

Table 3. Social media usage (T1–T3)^a in 12 Western European countries (YouTube and Twitter).

Country ^b	YouTube, n (%)			P value		Twitter, n (%)			P value	
	T1	T2	T3	T1 vs T2 vs T3 ^c	T1 vs T3 ^d	T1	T2	T3	T1 vs T2 vs T3 ^c	T1 vs T3 ^d
NL (n = 88)	9 (10%)	23 (26%)	33 (38%)	<.001	<.001	4 (5%)	27 (31%)	49 (56%)	<.001	<.001
BE (n = 91)	1 (1%)	4 (4%)	5 (5%)	.04	.14	0	2 (2%)	6 (7%)	.009	.03
LU (n = 6)	0.0	1 (17%)	1 (17%)	.37	1	0	0	0	ND ^e	ND
DE (n = 347)	3 (1%)	20 (6%)	52 (15%)	<.001	<.000	2 (1%)	9 (3%)	23 (7%)	<.001	<.001
AT (n = 25)	0	3 (12%)	3 (12%)	.05	.25	0	0	0	ND	ND
CH (n = 41)	0	2 (5%)	5 (12%)	.02	.06	0	1 (2%)	1 (2%)	.37	1
UK (n = 175)	6 (3%)	37 (21%)	62 (35%)	<.001	<.001	4 (2%)	42 (24%)	68 (39%)	<.001	<.001
IR (n = 28)	0	0	0	ND	ND	0	0	1 (4%)	.37	1
NO (n = 17)	0.0	2 (12%)	3 (18%)	.1	1	0	2 (12%)	8 (47%)	.002	.008
SE (n = 22)	0	5 (23%)	5 (23%)	.007	1	0	2 (9%)	2 (9%)	.14	.5
FI (n = 9)	0	0	0	ND	ND	0	0	0	ND	ND
DK (n = 24)	0.0	2 (8%)	3 (13%)	.1	.25	0	0	0	ND	ND
All (n = 873)	19 (2%)	99 (11%)	172 (19.7%)	<.001	<.001	10 (1%)	85 (10%)	158 (18.1%)	<.001	<.001

^a T1 = April to August 2009, T2 = August to December 2010, T3 = April to July 2011.

^b NL = the Netherlands, BE = Belgium, LU = Luxembourg, DE = Germany, AT = Austria, CH = Switzerland, UK = United Kingdom, IR = Ireland, NO = Norway, SE = Sweden, FI = Finland, DK = Denmark.

^c Cochran Q test (df = 2).

^d McNemar test (df = 2).

^e No data.

Table 4. Social media usage (T1–T3)^a in 12 Western European countries (Facebook, blogs, and LinkedIn).

Country ^b	Facebook, n (%)			<i>P</i> value	Blog, n (%)			<i>P</i> value		LinkedIn, n (%)			<i>P</i> value
	T1	T2	T3		T1	T2	T3	T1 vs T2 vs T3 ^d	T1 vs T3 ^c	T1	T2	T3	
NL (n = 88)	ND ^e	0	13 (15%)	<.001	2 (2%)	5 (6%)	4 (5%)	.1	.5	ND	48 (55%)	71 (81%)	<.001
BE (n = 91)	ND	20 (22%)	62 (68%)	<.001	2 (2%)	2 (2%)	2 (2%)	1	1	ND	20 (22%)	41 (45%)	<.001
LU (n = 6)	ND	0	3 (50%)	.25	0	0	0	ND	ND	ND	0	2 (33%)	.5
DE (n = 347)	ND	26 (8%)	232 (66.9%)	<.001	0	0	1 (1%)	.37	1	ND	6 (2%)	10 (3%)	.22
AT (n = 25)	ND	1 (4%)	21 (84%)	<.001	0	0	0	ND	ND	ND	1 (4%)	3 (12%)	.5
CH (n = 41)	ND	4 (10%)	15 (37%)	.001	0	0	0	ND	ND	ND	5 (12%)	9 (22%)	.13
UK (n = 175)	ND	31 (18%)	163 (93.1%)	<.001	0	10 (6%)	12 (7%)	<.001	<.001	ND	71 (41%)	97 (55%)	<.001
IR (n = 28)	ND	0	23 (82%)	<.001	0	0	0	ND	ND	ND	0	3 (11%)	.25
NO (n = 17)	ND	2 (12%)	15 (88%)	<.001	0	0	1 (6%)	.37	1	ND	8 (47%)	13 (76%)	.06
SE (n = 22)	ND	0	10 (45%)	<.001	3 (14%)	3 (14%)	2 (9%)	.37	1	ND	15 (68%)	17 (77%)	.5
FI (n = 9)	ND	0.0	7 (78%)	.02	0	0	0	ND	ND	ND	0	1 (11%)	1
DK (n = 24)	ND	1 (4%)	21 (88%)	<.001	0	0	1 (4%)	.37	1	ND	5 (21%)	11 (46%)	.03
All (n = 873)	ND	85 (10%)	585 (67.0%)	<.001	7 (1%)	20 (2%)	23 (3%)	<.001	<.001	ND	179 (20.5%)	278 (31.8%)	<.001

^a T1 = April to August 2009, T2 = August to December 2010, T3 = April to July 2011.

^b NL = the Netherlands, BE = Belgium, LU = Luxembourg, DE = Germany, AT = Austria, CH = Switzerland, UK = United Kingdom, IR = Ireland, NO = Norway, SE = Sweden, FI = Finland, DK = Denmark.

^c McNemar test (df = 2).

^d Cochran Q test (df = 2).

^e No data.

Table 5. YouTube videos, views, and Twitter followers at T2 and T3^a.

Country ^b	YouTube videos per account, median (IQR ^c)		<i>P</i> value T2 vs T3 ^d	YouTube views per account, median (IQR)		<i>P</i> value T2 vs T3 ^d	Twitter followers per account, median (IQR)		<i>P</i> value T2 vs T3 ^d
	T2	T3		T2	T3		T2	T3	
NL (n = 88)	5 (2–20)	9 (5–26)	.03	839 (221–1721)	4828 (976–12022)	<.001	119 (48–235)	336 (150–748)	<.001
BE (n = 91)	3 (2–3)	7 (3–9)	.18	241 (145–241)	6648 (3332–13241)	.04	175	127 (41–232)	ND ^e
LU (n = 6)	5 (a)	4 (a)	ND	141	244	ND	0	0	ND
DE (n = 347)	2 (2–17)	6 (3–19)	.01	1809 (737–27,823)	1920 (382–11366)	.001	51 (27–76)	90 (30–309)	.18
AT (n = 25)	20 (10–22)	32 (18–58)	.11	10,930 (5465–12,755)	26,251 (19,855–29,692)	.18	0	0	ND
CH (n = 41)	2 (2–3)	6 (3–16)	1	3 (a)	3717 (2003–3853)	ND	19	63	ND
UK (n = 175)	5 (2–8)	7 (4–16)	.004	256 (137–1436)	2372 (880–7313)	<.0001	311 (135–625)	464 (145–1019)	<.001
IR (n = 28)	0	0	ND	0	0	ND	0	44 (a)	ND
NO (n = 17)	7 (5–8)	4 (3–8)	.32	2962 (2700–3223)	5250 (5200–7082)	.18	57 (30–83)	200 (65–370)	.18
SE (n = 22)	13 (7–16)	12 (4–12)	.13	560 (458–7199)	3146 (1892–12029)	.35	84 (75–92)	142 (116–169)	.18
FI (n = 9)	0	0	ND	0	0	ND	0	0	ND
DK (n = 24)	1 (1–2)	3 (3–3)	ND	101 (51–152)	120.0 (71–168)	ND	0	0	ND
All (n = 873)	4 (2–13)	7 (3–16)	<.001	575 (190–2444)	3074 (724–10110)	<.001	204 (74–579)	271 (85–724)	<.001

^a T1 = April to August 2009, T3 = April to July 2011.

^b NL = the Netherlands, BE = Belgium, LU = Luxembourg, DE = Germany, AT = Austria, CH = Switzerland, UK = United Kingdom, IR = Ireland, NO = Norway, SE = Sweden, FI = Finland, DK = Denmark.

^c Interquartile range.

^d Wilcoxon signed rank test.

^e No data.

YouTube

YouTube accounts were found in 10 countries (Table 3). At T3, we found significant differences in the percentage of YouTube usage ($\chi^2_{11} = 73.9$, $P < .001$). The Netherlands (38%, $n = 33$) and the United Kingdom (35%, $n = 62$) had the highest percentage of hospitals with a YouTube account. During the research period, the percentage of YouTube accounts increased significantly (Table 3). The median number of videos per YouTube account at T3 was 7 (Table 5).

Twitter

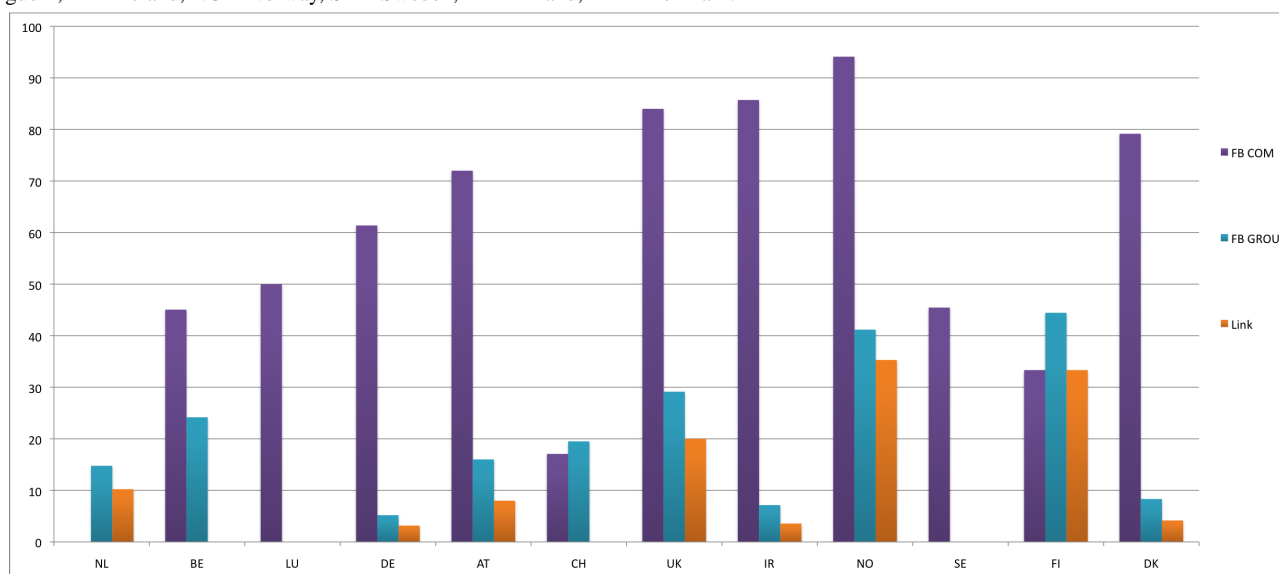
Twitter accounts were found in 8 of 12 countries (Table 3), with significant differences between countries ($\chi^2_{11} = 209.2$, $P < .001$) at T3. The Netherlands (56%, $n = 49$), the United Kingdom (39%, $n = 68$), and Norway (47%, $n = 8$) had the highest

percentages of hospitals with a Twitter account. The median number of followers for all countries at T3 was 271 (Table 5). We identified 1 hospital with 3300 followers.

Facebook

Facebook accounts were found in all countries, ranging from 15% ($n = 13$) in the Netherlands to 93.1% ($n = 163$) in the United Kingdom (Table 4). At T3, there was a significant difference between all countries in the percentage of Facebook usage ($\chi^2_{11} = 202.1$, $P < .001$). Facebook usage increased significantly in 11 countries. Two types of Facebook accounts were found: company profiles and group pages (Figure 1). The number of Facebook group pages was lower, ranging from 0% in Luxembourg to over 40% in Finland and Norway. Apart from 2 countries (Norway and Finland), having a Facebook page accessible through the hospital's website was an exception (Figure 1).

Figure 1. Percentage of Facebook company profiles (FB COM), group pages (FB GROUP), and links (Link) to a Facebook account on hospital websites at T3 (April to July 2011). NL = the Netherlands, BE = Belgium, LU = Luxembourg, DE = Germany, AT = Austria, CH = Switzerland, UK = United Kingdom, IR = Ireland, NO = Norway, SE = Sweden, FI = Finland, DK = Denmark.



Blogs

Blogs were found in 7 of the 12 countries, ranging from 1% ($n = 1$) in Germany to 9% ($n = 2$) in Sweden (Table 4). We found blogs less frequently than the other types of social media. The percentages of blogs differed significantly between countries ($\chi^2_{11} = 28.5$, $P = .003$).

LinkedIn

We measured LinkedIn during two periods (T2 and T3). We found significantly increased usage in 4 countries. At T3, the percentage of LinkedIn accounts ranged from 3% ($n = 10$) in Germany to 81% ($n = 71$) in the Netherlands (Table 4), and the percentages were significantly different ($\chi^2_{11} = 336.4$, $P < .001$). Of all 873 hospitals, we found 1 hospital with a link to their LinkedIn profile on their website.

Discussion

In this longitudinal study we explored the use of social media by 873 hospitals in 12 Western European countries. The use of social media increased in all of the countries, especially YouTube (from 2% to 19.7%), LinkedIn (20.5% to 31.8%), and Facebook (10% to 67.0%). This increased use of social media has been confirmed by other studies [14]. Interestingly, the use of Twitter, Facebook, and YouTube in Europe appeared to be higher than in the United States [15].

There are notable differences between the 12 countries. The use of Twitter was especially popular in the United Kingdom, the Netherlands, and Norway. At the third measurement, almost half of all hospitals in the Netherlands and in Norway were on Twitter. YouTube was used by 35% of the hospitals in the United Kingdom and 38% in the Netherlands, whereas the use of YouTube varied from 0% to 23% in all other countries. There are several possible reasons for the differences between countries that we found. First, the use of social media could be related to the Internet penetration in a specific country. However,

the differences in broadband penetration in Europe are small [21]. Second, there may be an influence of local or country-specific social media. An example is Hyves, which was, until recently, the most popular social network in the Netherlands, with more than 11 million members [22]. This could explain why Facebook was less popular in the Netherlands than in other countries. It is difficult to predict the popularity or influence of other social media. Online sources show that Facebook, when Hyves is excluded, was the most popular social media network in all other countries during the research period [19,20].

The activity of hospitals on social media increased during the research period, as the number of videos and viewers of YouTube channels, and of Twitter followers increased. Furthermore, the increased usage of LinkedIn was notable in the Netherlands and the United Kingdom at the third measurement. Hospitals in these countries seem to be aware of the benefits of recruiting personnel that LinkedIn offers. However, the observation that only 5% ($n = 48$) of all 873 hospitals had a link to their YouTube channel and 10% ($n = 90$) had a link to their Twitter feed on their website indicates that hospitals are not using the full potential of all types of social media yet. Based on this study, we cannot say anything about the content of videos, tweets, and messages. However, our data show that an ever-increasing number of users are watching the videos and reading the tweets.

Since Western European hospitals have become aware of social media and increasingly use it, we foresee great opportunities to improve health care and to stimulate participatory health care. Various studies have described improvements that social media could offer to health care, such as greater transparency, openness, and communication, and improved patient support and knowledge translation [4,10]. Therefore, research should be focused on describing best practices, which may help speed up implementation of social media. Furthermore, it would be worthwhile to identify for what purposes hospitals use social media and to what extent social media improve participatory

health care. For a complete overview, future research should also focus on the challenges and risks of using social media, such as legal constraints, fraud, and budget constraints. These topics are also important research subjects in the light of the discussion about desirability of social media usage by health care professionals.

Our study has some limitations that need to be discussed. In a few cases, we experienced difficulties determining whether a social network was official (was initiated and maintained by the hospital itself). However, we gave hospitals the opportunity to correct their data. Another aspect is the differences between health care systems in the included countries. We found that in a few countries, some hospital organizations included more than 1 hospital. Since we counted these organizations as 1 hospital, our data do not reflect the results of individual hospitals in every country.

Another aspect is that we measured Facebook and LinkedIn only at T2 and T3. It would have been interesting to see the results for T1. However, at the start of the project, we were not aware of hospitals using Facebook or LinkedIn. Since Facebook and LinkedIn became increasingly popular in 2009 and 2010, we decided to include them in the search we conducted in this study.

Awareness and use of social media is growing in Western European hospitals. Social media usage differs significantly between countries. Except for the Netherlands and the United Kingdom, the group of hospitals that are using social media remains small. Usage of LinkedIn for recruitment of personnel shows that hospitals are aware of the potential of social media. Future research is needed to investigate how social media lead to improved health care.

Conflicts of Interest

None declared.

References

1. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. *Bus Horiz* 2010;53(1):59-68. [doi: [10.1016/j.bushor.2009.09.003](https://doi.org/10.1016/j.bushor.2009.09.003)]
2. Wikipedia. 2011 Oct 11. Social Media URL: http://en.wikipedia.org/wiki/Social_media [accessed 2011-10-13] [WebCite Cache ID 62PR0BqUP]
3. Eurostat Press Office. European Commission. 2010 Dec 14. Internet Access and Use in 2010: 80% of Young Internet Users in the EU27 Active on Social Media URL: http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/4-14122010-BP/EN/4-14122010-BP-EN.PDF [accessed 2011-10-13] [WebCite Cache ID 62PS2QwZi]
4. Bacigalupe G. Is there a role for social technologies in collaborative healthcare? *Fam Syst Health* 2011 Mar;29(1):1-14. [doi: [10.1037/a0022093](https://doi.org/10.1037/a0022093)] [Medline: [21417520](https://pubmed.ncbi.nlm.nih.gov/21417520/)]
5. Facebook. 2012. Statistics URL: <http://newsroom.fb.com/content/default.aspx?NewsAreaId=22> [accessed 2012-04-10] [WebCite Cache ID 66pOLrDLk]
6. US Census Bureau. 2011 Oct 13. US and World Population Clocks URL: <http://www.census.gov/main/www/popclock.html> [accessed 2011-10-13] [WebCite Cache ID 62PSd8ju7]
7. Barnes NG, Mattson E. University of Massachusetts Dartmouth Center for Marketing Research. 2010. The Fortune 500 and Social Media: A Longitudinal Study of Blogging and Twitter Usage by America's Largest Companies URL: http://dl.dropbox.com/u/395938/Barnes_Mattson_2010.pdf [accessed 2012-04-17] [WebCite Cache ID 66zGtnKL8]
8. Van De Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. *J Med Internet Res* 2010;12(2):e18 [FREE Full text] [doi: [10.2196/jmir.1350](https://doi.org/10.2196/jmir.1350)] [Medline: [20542857](https://pubmed.ncbi.nlm.nih.gov/20542857/)]
9. Thinkinsights. Google/OTX. 2009 Dec. Health Consumer Study: The Role of Digital in Patients' Healthcare Action & Decisions URL: http://www.gstatic.com/ads/research/en/2009_HealthConsumerStudy.pdf [accessed 2012-04-13] [WebCite Cache ID 66tWju0o7]
10. The Change Foundation. 2011 Jun. Using Social Media to Improve Healthcare Quality URL: <http://www.changefoundation.ca/docs/socialmediatoolkit.pdf> [accessed 2011-11-09] [WebCite Cache ID 633zuJfkb]
11. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000 Oct;51(7):1087-1110. [Medline: [11005395](https://pubmed.ncbi.nlm.nih.gov/11005395/)]
12. Thielst CB. Using social media to engage patients: many tools exist to connect, communicate and build loyalty. *Healthc Exec*. 2011. (3) p. 66-70 URL: <http://thielst.typepad.com/files/satisfying-your-customers-with-social-media.pdf> [WebCite Cache ID 67Ff3BQIs]
13. Gamble KH. Just a tweet away. *Healthc Inform* 2009 May;26(5):30, 32, 34 passim. [Medline: [19514237](https://pubmed.ncbi.nlm.nih.gov/19514237/)]
14. Hawker MD. Social networking in the National Health Service in England: a quantitative analysis of the online identities of 152 primary care trusts. *Stud Health Technol Inform* 2010;160(Pt 1):356-360. [Medline: [20841708](https://pubmed.ncbi.nlm.nih.gov/20841708/)]
15. Bennett E. Found In Cache. 2010 Oct 19. Hospital Social Network List URL: <http://ebennett.org/hsnl/> [accessed 2011-11-09] [WebCite Cache ID 633zdEW1V]
16. American Hospital Association. 2011. Fast Facts on US Hospitals. Sep 14 URL: <http://www.aha.org/research/rc/stat-studies/fast-facts.shtml> [accessed 2011-11-09] [WebCite Cache ID 633zhyujA]

17. Wikipedia. 2011 Aug 20. Lists of Hospitals in Europe URL: http://en.wikipedia.org/wiki/Lists_of_hospitals_in_Europe [accessed 2012-01-14] [WebCite Cache ID 64gU1NGfd]
18. MediLexicon International Ltd. 2012. Worldwide Hospitals Directory URL: <http://www.hospitalsworldwide.com/> [accessed 2012-01-14] [WebCite Cache ID 64gU9UaUN]
19. Vincos Blog. 2012. World Map of Social Networks URL: <http://vincos.it/world-map-of-social-networks/> [accessed 2012-01-14] [WebCite Cache ID 64gUCPvux]
20. Edelman.com. 2010 Nov 19. Social Media Mapping in Europe URL: http://www.slideshare.net/EdelmanDigital/social-media-mapping-in-europe?from=ss_embed [accessed 2012-01-14] [WebCite Cache ID 64gUNFNQH]
21. Miniwatts Marketing Group. 2011 Aug 9. Internet and Facebook usage in Europe URL: <http://www.internetworldstats.com/stats4.htm> [accessed 2011-11-09] [WebCite Cache ID 633zlljWe]
22. Telegraaf Media Groep. 2012. Hyves URL: <http://www.hyves.nl/about/press/> [accessed 2012-02-24] [WebCite Cache ID 65glvugo0]

Edited by B Bewick; submitted 09.11.11; peer-reviewed by B Meskó, T Biswas; comments to author 07.12.11; revised version received 01.03.12; accepted 07.03.12; published 01.05.12.

Please cite as:

Van de Belt TH, Berben SAA, Samsom M, Engelen LJLPG, Schoonhoven L
Use of Social Media by Western European Hospitals: Longitudinal Study

J Med Internet Res 2012;14(3):e61

URL: <http://www.jmir.org/2012/3/e61/>

doi: [10.2196/jmir.1992](https://doi.org/10.2196/jmir.1992)

PMID: [22549016](https://pubmed.ncbi.nlm.nih.gov/22549016/)

©Tom H Van de Belt, Sivera AA Berben, Melvin Samsom, Lucien JLPG Engelen, Lisette Schoonhoven. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers

Liesje Donkin¹, BSc, MSc; Nick Glozier^{1,2}, MBBS, FRANZCP, PhD

¹Brain and Mind Institute, The University of Sydney, Camperdown, Australia

²Disciplines of Psychiatry and Sleep Medicine, Sydney Medical School, The University of Sydney, Camperdown, Australia

Corresponding Author:

Liesje Donkin, BSc, MSc

Brain and Mind Institute

The University of Sydney

Level 3

94 Mallet Street

Camperdown, NSW 2050

Australia

Phone: 61 293510520

Fax: 61 293510855

Email: Liesje.Donkin@sydney.edu.au

Abstract

Background: Many users of Internet interventions do not persist with the full treatment program. As persistence may influence outcomes of such interventions, being able to maximize persistence is vital. However, while studies have begun to explore the predictors of dropout in Internet interventions, few have explored reasons why users persist with the programs, which may not just be the converse of the reasons for dropout.

Objective: To answer the question of what influences persistence with online interventions.

Methods: We interviewed participants in the Cardiovascular Risk E-couch Depression Outcome (CREDO), a trial evaluating the efficacy of an eHealth intervention (e-couch) in treating depressive symptoms in those with comorbid depression and cardiovascular risk factors. Interviews were semistructured in nature and were analyzed using a grounded theory approach. Interview numbers were curtailed ($n = 12$) after theoretical saturation.

Results: All participants reported substantial barriers to completing the program including time constraints, competing priorities, anxiety about spending time on the computer, and perception of limited worth of the intervention. Participants who persisted with the trial reported intrinsic motivations such as personal values about task completion and sense of control, and recognized external motivators that aided the development of habits and identified personal benefits attributable to the program.

Conclusions: Online interventions may benefit from content that enhances the intrinsic motivations such as a having sense of control and being able to identify with the program, and by increasing the relative value of the program in order to enhance persistence. Persistence within a trial setting appears modifiable through explicit messages regarding supporting others. In terms of motivators, the use of a hook to engage participants who are starting the intervention due to curiosity and the use of reminder systems to prompt participants may also improve persistence. The worth of such additions should be evaluated using adherence and outcomes metrics.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR): ACTRN12610000085077; <http://www.anzctr.org.au/ACTRN12610000085077.aspx> (Archived by WebCite at <http://www.webcitation.org/68MtyPO3w>)

(*J Med Internet Res* 2012;14(3):e91) doi:[10.2196/jmir.2100](https://doi.org/10.2196/jmir.2100)

KEYWORDS

Adherence; persistence; eHealth; online interventions; motivation; barriers

Introduction

As acceptability of the Internet and Web-based programs increases [1-5], so too does the volume of Web-based interventions for mental health and health behaviors [6-10]. Such interventions are not necessarily standalone interventions, but may be used in a stepped-care model [11,12], as an adjunct to face-to-face therapy to enhance the availability of access to mental health services [4,13], or as an intervention supported by a health care worker [14,15].

Increasing interest in online interventions or e-therapy has largely been driven by the belief that delivering interventions over the Internet overcomes some of the barriers of traditional face-to-face therapy. Such issues include geographical access, time restrictions, cost of accessing services, and reluctance to seek help [16-18]. The Internet as a mode of delivery may overcome these barriers through being available where and when the client needs the intervention [1], and by promoting help seeking and disclosure through anonymity [19-21]. Additionally, some studies have found that online interventions are less affected by behavioral avoidance of session attendance [22], are more cost effective [23], and may reach people whose symptoms are more severe than the symptoms of those who present to a clinic [1]. Importantly, some research participants have also indicated that they perceived online interventions as a valid [24] and credible [4] form of treatment with high levels of acceptability.

While online interventions have been shown to be effective [25-27], little is known about the degree to which the users' engagement in, persistence with, and use of the program (or program adherence) influences the intervention outcomes. In e-therapy research trials, the term dropout is often used in place of persistence and refers to when the individual does not complete or persist with the intervention, thereby not completing the required modules or assessments [28]. However, a participant can fail to persist with the program, yet not drop out from the trial by continuing to provide outcome data. This can affect the ability of results to be generalized [29,30] and undermine the outcome of the trial. To date, efficacy trials have generally reported, when mentioned, good to excellent levels of persistence, often due to the intensive preparations, assistance, and follow-up of participants [28]. However, many authors have noted that dropout rates are high in open access trials [28,31-33], with only a small proportion persisting with the trial and associated follow-ups. For example, Christensen et al reported that only 0.5% of spontaneous users completed an open access online depression program [34], with similar figures of 1% in an open access trial for anxiety [35]. However, the dropout rates may vary according to the intervention features, with clinician-guided interventions having less dropout than interventions that are fully automated [35,36] and closed trials having better persistence (ranging from 47% to 99% [28]) than open access trials [34,37], purportedly due to the follow-up by the research team [32]. Alternatively, the low persistence rates in open access trials may be an artifact of "window shopping," where potential users register for the program to determine what the program is like. Such an approach leads to the user completing 1 or a few modules in order to decide whether to

continue with the program. The intensive preparation, informed consent, and screening processes within closed trials make such window shopping less likely, given the upfront investment and awareness in the program before commencement.

Research into program engagement, particularly the reasons for nonadherence and poor persistence, is sparse despite the thinking that the degree of engagement in self-directed or self-help online therapy is as important as engagement in face-to-face therapies [10,38,39] for optimizing outcome. What data there are indicate that users frequently cite forgetfulness and a lack of time as reasons for poor adherence [40,41] or persistence.

Maximizing a participant's motivations to persist seems a self-evident goal of an e-therapy program [2,42], even though the moderating and mediating factors in online intervention effectiveness are not well understood [43]. Theoretical models have been developed to explain adherence and persistence on the basis of factors determined a priori by researchers in quantitative studies; drawn from medication adherence literature, health behavior change models, and technology studies; or based on the quantitative findings from online interventions. Only two studies have qualitatively explored the user perspectives and motivations that may influence adherence and persistence with asynchronous (where feedback or communication does not occur in real time but has a time delay) or automated e-therapy interventions. Gerhards et al completed a qualitative study using 18 semistructured interviews [44], concluding that three types of factors influenced adherence or persistence to, and the perceived effectiveness of, the trial: technical factors, such as knowledge and ability to access a computer; social factors, such as desire for people contact, levels of motivation, and belief that the program was applicable to the individual's situation; and research-specific factors that may influence symptoms and symptom reporting. In a different approach, Bendelin et al classified people who participated in a guided self-help online intervention into readers, strivers, and doers based on interviews with 12 strategically selected participants [45]. Overall, motivation to persist with the intervention was influenced by participants' perception of support they received and required, and by their perception of improvement.

While these two studies provide an understanding of some of the factors that influence adherence and the translation of online to offline behavior in their perspective populations, no clear model exists to explain the role of motivation in influencing persistence in online interventions. Grounded theory is a form of systematic qualitative inquiry that allows the generation of theory from data. This methodology is useful for the development of new theories and ideas in areas where theory does not readily exist, such as the area of motivation in online interventions and e-therapy. This study took a grounded theory approach to address the question of what influences persistence with online interventions in order to propose a model to aid program development.

Methods

We conducted face-to-face and telephone interviews with participants in the Cardiovascular Risk E-couch Depression Outcome (CREDO) study, which evaluated the efficacy of an

eHealth intervention (e-couch) in treating depressive symptoms in those with comorbid depression and cardiovascular risk factors [46].

Participants, recruited from those engaged in the study between the 6- and 12-month follow-up stages, were invited by email to participate in the qualitative study. Participants who were less than 2 weeks away from an assessment were excluded, so as not to add burden to their participation for the final survey. We sent 80 email invitations to participants with the aim of recruiting approximately 10% of these. Participants were invited only once. Participants were selected at random from those eligible using a random number generator. Recruitment occurred in three waves and continued until theoretical saturation (where no new insights emerged from the data) [47] was reached. Interviews were offered without knowledge of the study arm and were not stratified. Participants were not remunerated for their participation in either the CREDO trial or this qualitative substudy.

Intervention

E-couch is a 12-week online cognitive behavioral therapy intervention for depression. The program contains elements pertaining to psychoeducation, activity scheduling, thought challenging, problem solving, and interpersonal therapy. The program is fully automated with a new module opening each week. During the trial, participants received a reminder email 3–4 days after the module had opened. They also received a scripted reminder, but not therapeutically assisting, telephone call from a research assistant if they had still not completed the module 1 week after it had opened.

Ethics Approval

We obtained written informed consent from all participants. Ethics approval for the CREDO trial and the qualitative substudy was obtained from the University of Sydney Human Research Ethics Committee. The University of New South Wales Human Research Ethics Committee provided ethics approval for the 45 and Up Study.

Study Sample

Participants in the study were residents of New South Wales, Australia, who had enrolled previously in the 45 and Up Study [48] and, subsequently, the CREDO trial [46]. Given this, all participants were aged over 45 years, had an active email address, had indicated a self-reported history or high risk of cardiovascular disease, and had at least a moderate level of depressive symptoms at two time points.

Data Gathering

We conducted 10 telephone and two face-to-face interviews between August and December 2011 at a location and time of the interviewee's convenience. The interviews were conducted and analyzed using a grounded theory approach, where themes and theories were allowed to emerge without using preconceived hypotheses and ideas [49]. An initial semistructured interview guide was developed with new questions being reformulated during the transcription and data analysis and added to the interview schedule in an iterative fashion. Therefore, the interview schedule evolved over the course of the research

project. Interviews lasted between 46 and 68 minutes, with all interviews revolving around one key question: "What kept you using the online program after you had begun?" Theoretical saturation appeared by the ninth interview, with a further three interviews being completed to ensure that this was the case.

After we received informed consent from participants, we audio recorded interviews using an electronic dictating machine.

Data Analysis

On completion of interviews, the interviewer transcribed audio files verbatim and then checked the transcriptions against the audio recordings. Interview transcripts were analyzed and coded as soon as possible after each interview. Handwritten field notes and impressions of the interview were also used to inform the analysis and generation of memos. The lead researcher analyzed typed transcripts of the interview using codes developed from written comments on the transcription. These comments were analyzed using a grounded theory approach emphasizing the use of iterative techniques, whereby data and emergent theory are constantly compared [49]. During the initial coding process, the lead researcher coded the transcripts to identify ideas in the data. This was followed by focused coding, where a central set of codes was pursued based on the prevalence of the initial codes and those considered a priori by the researcher. Theoretical coding was then used to link the codes to each other. Conceptual memos were written from the focused codes to help develop an understanding of the codes and how these related to the data and other codes. No coding or analysis software was necessary for this analysis.

Results

In total, two central themes of barriers and motivators emerged, each with several subthemes. All participants reported substantial barriers that decreased their motivation to continue with the program. These included time constraints and competing priorities, technology fatigue from spending the day on the computer at work, anxiety about spending time on the computer away from other demands of day-to-day life, and perception of limited worth of the intervention.

Participants who persisted with the trial identified intrinsic motivations and extrinsically motivated strategies to overcome the barriers that they faced in persisting with the trial. These included developing habits, recognizing personal values about completion, and identifying the benefits for others if the benefits for themselves were not immediately obvious.

Barriers

Personal Factors that Decreased Intrinsic Motivation to Participate

Forgetfulness or a Lack of Awareness That the Program Needs to be Completed

The competing demands for time and need to prioritize also appeared to result in many participants "forgetting" to complete their module and associated activities on a weekly basis. Many participants alluded to relying on weekly reminders that a module had opened, rather than initiating seeking out the new

module themselves. Forgetfulness appeared to be particularly prevalent when the participants postponed completing a module until a later time. Many stated that they would put the module off, intending to complete it when they would be able to focus on the intervention without competing demands, but often forgot to return to it. They often forgot when other demands on their time arose or when they lacked regular computer time, resulting in a lack of visible reminders that the program was open.

It was good, but I did especially appreciate the reminder because sometimes it came through at a busy time, um, I didn't mean to forget about it but it happens. And I was thankful for the reminder. [Interviewee #11]

Mood/Anxiety

Mood and anxiety were also often cited as barriers to engagement. When anxiety was high, participants found it difficult to find the time or concentrate on the program. When mood was low, participants lacked motivation to complete the program, often feeling overwhelmed by its demands. This resulted in participants focusing on other more manageable tasks, which allowed them to feel that they had completed at least some tasks throughout the day. Tasks of daily living were often prioritized ahead of the program, even when participants believed that they would benefit from taking the time to complete the program and objectively had the time.

...I get anxious, and then I begin to think to myself that "I've got to do this" and "I've got to do that," and "I can't do my [program] now, I'll do it later," and really there is nothing that can't wait. Nothing at all. I have a set routine. I'm retired. But I get myself into such a state of anxiety that I can't relax and do my [program]. So I leave it and go and do my silly little things such as taking my dog for a walk and doing my shopping. All sorts of, you know, mundane things that are not important, well they are important, but I could give myself the time to relax and do it... [Interviewee #1]

Factors That Reduced Engagement With the Program

Program Seen to be a Poor Fit to the Individual

Several participants reported frustration with the program's lack of personalization. This was particularly reinforced by frustration with the standardized questionnaires embedded in the program and trial. Participants frequently cited that the answers were unlikely to reflect their true state, as they did not believe that they fit within the categories provided. This resulted in them feeling frustrated and wanting to stop completing the questionnaires. Many indicated that the capacity to provide additional information in terms of a free-text box would be beneficial and would enhance their desire to persist with the program.

I found the questions strange. The question that said "Do you do what your doctor tells you?," I'm going "I'm not 4,...how am I going to respond to this?" And in the end I throw my arms in despair and go "Urgh! I don't know, I don't know the answer!" and put it in

the appropriate box...I found that when I clicked on the "Do I take x, y, z medicine" and put "Yes" and the window came up and gave me space to write in it. I want more of those windows. [Interviewee #4]

Other participants were concerned that their answers were being misinterpreted, which may have influenced their answers and therefore the time taken to answer the questionnaires or complete program activities. This was reflected in the number of emails participants sent, unsolicited, to the project team to contextualize their answers to the questionnaires.

The perceived unidirectional relationship between the user and the program also led participants to consider that how they engaged in the program was irrelevant and decreased their motivation to continue. They felt that the computer offered little in terms of support, interaction, and feedback. Participants frequently reported thoughts of giving up on the questionnaires and the trial due to feeling that their information would be misrepresented and therefore not useful to the trial team.

Because you need a bit of feedback...It's a bit of an empty system, because...you know...you put out a few stories about your thoughts and experiences, but you know...nothing comes back. [Interviewee #1]

Although some participants reported that this unidirectional nature of the relationship was beneficial, for others, it reduced any obligation to log on to the system and use it. The sense of accountability was decreased, as there was no feedback or cost of engaging with the program. A lack of perception of the program being a therapeutic relationship and of someone relying on their attendance meant that several participants felt that they were able to put off completing the program. The freely available nature of the program, without cost, waiting lists, or time restrictions, meant that participants placed less value on completing the program when it was opened to them.

...it's a lot easier to ignore. Um...and truly and if you've made an appointment and you know you have to pay for it, you're going to turn up. [Interviewee #1]

Failure to Learn Anything New

Participants who felt they were not learning anything reported lower levels of motivation to persist with the program. Some had had previous psychological treatment for depression and felt that the program did not offer them anything substantially different.

...think I was a bit of a difficult ah...participant, because I know a lot about uh...um...emotional therapies. And...and...a lot about the theories, so...um... it was uh...it was a bit tedious to go through each of the programs and the theories and then the examples for me... [Interviewee #1]

...a lot of the stuff we did online was not new...and that frustrated me. [Interviewee #2]

Yeah, and the other thing you know, it's long-winded. I can, sure there are some people, but for people who have, maybe, who have been to university, who have a certain level of understanding of these problems or

a certain level of education um, it was too simplistic for me. [Interviewee #10]

Lack of Perception of the Program as a Therapeutic Relationship

Previous therapy experiences also meant that some had different expectations of the therapeutic relationship. The program therefore offered reminders and information, but not therapy as such. This meant that participants struggled to engage in the computer intervention and to see it as therapy.

Um...ah...because you have no feedback, they're like chalk and cheese really. Well, yes...it was educative, but I didn't feel like it was a therapy session... [Interviewee #1]

Those participants who perceived little personal benefit from the program reported persisting due to the perception of obligation to the researcher or the belief that their input would help the wider community. They appeared to hold values around completion and contributing to society that allowed them to continue to persist with the program.

Motivating Factors That Enhanced Persistence and Adherence to the Program

Participants reported several strategies and motivations that allowed them to complete the program on a (semi)regular basis. These strategies were rarely linked to a specific intervention barrier but were rather seen as intrinsic traits or behaviors that allowed them to continue engaging with the program.

Noticing an Improvement

The perception of receiving a benefit from the program was considered to be the primary reason many participants persisted with the intervention. Participants reported experiencing benefits of the program beyond a sense of accomplishment and looking forward to completing the next part of the program. These participants frequently spoke about the benefits of the program and how they were implementing changes from the program into their daily lives.

...if I thought to myself this is useless I'm not getting anywhere. I possibly wouldn't finish it...I think that it was very, very interesting. There was a heck of a lot to learn from it, there was a lot I found, the reading, everything, I found I quite...I enjoyed it. There you go. [Interviewee #3]

Participants who saw personal value often used the language of the program. They were able to recall tasks they completed, appeared to have a better recall of the program, and continued to refer back to the program after they had completed it. They reported an overall sense of satisfaction that carried through each module and they appeared engaged in the program as a whole.

I felt satisfied at the end of it. That I'd done it, it was interesting and it was...I suppose it reinforced what I probably should be doing myself. It was easy to do and pleasant, and yeah I felt good about doing it... [Interviewee #5]

Feeling in Control

Several participants reported that they had dropped out of previous face-to-face therapeutic relationships due to the nature of the interactions. They reported feeling as though the therapist had been restrained by the amount of time available and that they were rushed by the time allocated to individual sessions. This resulted in their feeling unimportant and merely a patient rather than a person, and that their agenda was not as important as the agenda of the therapist.

...you could sit there and just actually take your time to do it. You know you could really think. Whereas when you're talking to somebody and you've got an hour or three-quarters of an hour or something, you really kind of you know...so time to me is the important thing. [Interviewee #3]

Being able to dictate the pace in the online intervention meant that they felt they got more benefit from it. The ability to choose which activities to complete and what areas to focus on for the activities resulted in many feeling as though their agendas were catered for in the program rather than that of the therapist. Due to this, they felt that they could be more honest with the program, and this led to greater reflection.

I think when I reflect on therapy you've got the lead up, going to [laughs] therapy, the anxiety beforehand, you know. You're going into the therapist's room, you're in someone else's domain, but it's like being in front of a stranger, you know, you're in front of a stranger and you feel you have to perform a little bit. [Laughs] whereas if you're by yourself, controlling it yourself, I think you can be more reflective about yourself. [Interviewee #11]

Many participants reported returning to the intervention several times. This allowed them to reflect on what they had learned. This appeared to facilitate the program being incorporated into everyday life. Those who returned to the program several times were more likely to find the program material novel and interesting. Several of these users often frequently returned to other trusted websites when they needed information. These participants appeared to view the program as a useful tool for themselves, rather than a research project where their contribution would be more valuable to others instead of themselves.

The ability to be flexible in the completion of the program also meant that participants were able to overcome time-related barriers, allowing them to complete modules in sections, when time allowed. Sometimes this appeared to be prompted by reminders that were part of the program.

I remember doing a long one and thinking, "Oh yeah! I missed one! I can go in and do that one again". It always good because you know you can go in and do it again. That's the first time I actually left the memory one and come back and do it again and I'm going "I'll do that later" [Interviewee #4]

The perception of flexibility and control of the program appeared to enhance engagement in the program and therefore improved motivation to persist with the trial.

Sense of Duty to Oneself

Many participants cited value in completing what they had started as a reason to continue the program despite being frustrated with it. The participants often contextualized these values as being a learning of their generation, with many indicating that these values had developed as they had grown up and had persisted into adulthood.

...because I had committed to doing something and so I did it. [Interviewee #2]

Due to having these values, they were going to complete the program regardless of how frustrating or tedious it was. They completed it because their sense of what they should and shouldn't do directed them. Holding such values meant that they did not consider stopping the program or dropping out from the study.

Sheer determination to see what's this one going to be about, curiosity perhaps and it was determination. I feel you know if a job's worth doing you do it as good as you can. There's no point in giving up in the middle of [a] thing you just keep on plodding on. [Interviewee #3]

On completing the entire program they described a sense of accomplishment coupled with relief, a sense of having done something that was difficult, but that they had overcome their difficulties. These participants placed value in hard work and were often scornful of those who dropped out, implying this was a weakness of character. They placed pride in completion where they saw the task as a chore rather than a personally useful tool or activity.

Something That Needs to be Done: Task Completion

Participants who integrated the program into their day-to-day routine appeared to be less likely to struggle to be adherent. They did not see the program as an activity to be reflected on, but rather as a task to be completed when it was available. These people did not report the frustrations with the program that those who saw completing the program due to a sense of duty did; rather, they saw it as a task to be done with little emotional connection to it.

I log in to my Internet every morning. I spend an hour or two hours, depending on what I have to do. And...um...and I'll always complete task as they come up. So with my emails, it's the same. I'll answer them on the same day or possibly the next day. [Interviewee #2]

This was more common in participants who logged on to the computer daily and who often referred to themselves in terms that implied a perception of being organized or task driven. Routine was often highly valued and the ability of the program to fit into this allowed it to become routine. These participants saw the program as something that needed to be done, rather than something that they wanted to do or enjoyed doing. They often described it as part of the daily chores or to-do list for that day. The development of habit and the perception of the modules as needing to be done is particularly interesting given that participants were able to opt out of the trial at any stage. Instead, they conceptualized the trial as medicine that needs to be taken.

Obligation to the Research and the Researchers

Many participants used words such as commitment and obligation to describe the things that made them remain engaged in the program. They described feeling a sense of commitment to the researcher or to the research. Many of these participants had been part of research teams in the past and therefore placed heavy emphasis on participating in research projects. Some were also aware of the numbers required to successfully complete a research project and aware of the impact of dropout on research outcomes.

Well, I mean, you really, you have to keep going. Um, when you start something you got to finish it, unless it drives you crazy. And it didn't drive me crazy. There were some days when I did it and I thought "God how much more of this?," but you made a commitment, so you followed through and you did it. And I mean if you don't finish it how can you help if you're only giving half information how can you help? And as I say the main criteria was perhaps I could help. [Interviewee #9]

I am thinking of the fact that I have committed. Commitment is all I can think of. And I am thinking that the person I have committed to is relying on my support. And um...it would be very unfair to let them down. I know that I am only one of many people, but if everybody would drop out...where would we be? [Interviewee #2]

When participants had engaged previously in programs that were not obviously linked to research, they were less likely to adhere to the program outline. This research-related engagement appeared to be driven by an obligation to others, rather than an obligation and perception of the benefits that they might have received themselves. Participants reported that when they were completing a program for only themselves, program use was more ad hoc, as they believed that they could complete the program at their own pace without affecting others. However, when engaged in research, there was more pressure to complete the program in a timely manner and in its entirety to ensure their contribution was able to benefit others.

For instance, with my [other program]. I drop in and I drop out depending on my other commitments because I know that nobody else is involved but myself. [Interviewee #2]

This motivating factor appears to be linked to the completion of research and may not enhance motivation in open access users of freely available online programs. The removal of an obligation to others when there is no research project attached to the program may lead to open access users not completing programs, not obtaining the full benefits of programs, or dismissing online programs as ineffective. Such experiences may also lead to participants dismissing psychological therapies as ineffective for their problem and may actually decrease help seeking.

Discussion

Overall, participants who had persisted with the e-therapy intervention reported multiple barriers and motivating factors in completing the program. It appeared that all participants initially engaged in the program with the hope that the program would be beneficial for themselves or for others. However, only half of the interviewed participants reported being able to complete their modules on a weekly basis. These participants appeared to have different barriers from those who completed the program less frequently, who often mentioned a lack of time and competing demands such as work and family. Motivating factors for less-adherent participants included believing that the research was important and feeling obligated to others. The more regularly adherent participants appeared to be older and reported more confusion around technology use but fewer issues with time demands and program frustrations. More-adherent participants were more likely to perceive the program as beneficial. These benefits consisted of mood improvements and a sense of getting things done, satisfaction at completing the program, and a sense of contributing to the wider community through research. However, over time the motivation to persist with the program decreased in many participants due to ongoing program frustration and the influence of personal factors such as mood issues and forgetfulness. To overcome these barriers, participants needed external motivators in the forms of reminders of their reasons for engagement and demonstrations of the benefits of completing the program (eg, changing symptom profiles). Consistent with previous research, these factors were often varied and multifaceted [42,44,45] but were incorporated into the program through the reminder emails and phone calls.

Several factors have emerged as themes about why people choose to engage and persist in an online intervention (see Figure 1.). Consistent with the medication adherence and persistence literature [50], this study found that people appear to have to perceive a benefit of the online intervention that is greater than the cost of participating. This benefit can be personal and, in the context of a trial, altruistic. Where there were time demands and pressures to complete work, participants often appeared to report a perceived benefit of completing the program as a factor that allowed them to persist with the program. This may be the result of the perception of having completed a to-do list task and through building a sense of mastery, which may be even more important when feeling overwhelmed and pressured from less-controllable areas of their lives. This indicates the importance of highlighting achievements within the program and reinforcing participation on an ongoing basis. A sense of obligation to oneself could also be manipulated through behavioral economic approaches. For instance, patients value and perceive gaining more benefit from health interventions for which they have paid [51] and may be more likely to use these as prescribed. To our knowledge no such approach has been tried in the e-therapy arena.

However, this study suggests two additional areas that influenced persistence other than personal benefits. The first of these, which is trial specific and may not be replicated in a nontrial or real-world setting, is that participants persisted with

the intervention and study because of beliefs that research is important, even when they felt little personal benefit from the program, or feelings of commitment to researchers. This is consistent with medical literature that has indicated that people who participate in research trials are more likely to view research more favorably [52] and are therefore more likely to participate. This indicates that highlighting the benefits of research in the early stages, such as recruitment, would help to initially engage people in the program and enhance their motivation to participate.

The second factor was the perception that through use of the program they may benefit the wider community. Consumers rarely appear to use interventions with the view that their behavior would affect the wider community, with the exception of vaccinations, which reduce the chance of disease for everyone through herd immunity, by decreasing exposure to pathogens. In terms of Internet-based interventions, users may be encouraged to reflect on the benefits of program completion. This might include consideration of the impact of their mood on others, the ability to provide feedback on the program so that it can be improved for others, and the importance of completing the program to determine for whom the program may be most useful, even if the participant is not finding it beneficial. While these are not necessarily trial factors per se, they may be more salient in a trial situation where participant information sheets and messages throughout the trial reinforce the experimental nature of the trial and the future benefits for others.

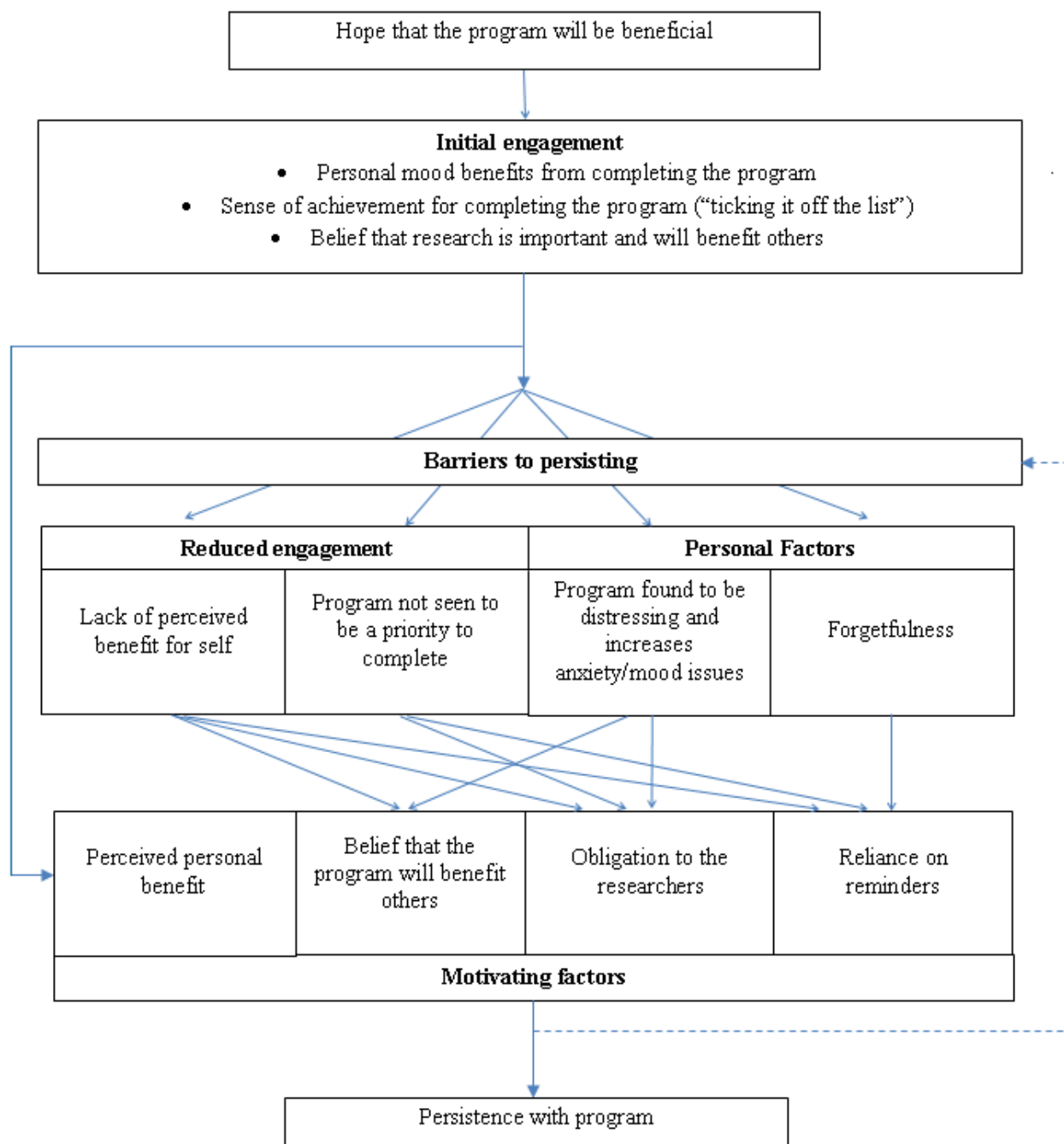
The influence of the research process on outcome has also been noted previously in online interventions [32,44]. Gerhards et al [44] noted that the research experience and the Hawthorne effect are likely to influence adherence and outcomes in online interventions, while Eysenbach noted that the reminders and telephone follow-up that were unique to their trial were also likely to have influenced persistence with the intervention [32]. This is consistent with our findings that reminders from the researchers influenced the participants' persistence. However, this study also found that many participants maintained implicit psychological contracts that resulted in their feeling obligated to complete the trial due to their perception of their relationship with the researchers. Developing an understanding of the perception of this therapeutic relationship and how this can be enhanced in online trials is important for facilitating uptake [42] and use of an intervention, and enhancing persistence with the intervention once the individual is engaged [53]. This is particularly important in online interventions without clinician support, where the users may feel less obligated to the program due to the absence of a clinician or the real person at the end of the telephone or email.

Certain predicted barriers and motivational factors were absent from the themes that emerged. We expected physical discomfort and health problems to be a barrier given the older comorbid nature of the population. Many participants reported physical discomfort to the trial team through email communication, and they also often mentioned physical health problems in the interviews, but participants did not talk about these as barriers to persistence. Participants considered psychological barriers and motivators to be of greater importance than physical ones.

This may reflect either that the intervention had a primary psychological target or that the Internet-based method of delivery may reduce travel-associated physical barriers. It may also represent a bias in the questions we used or in participants' placing greater emphasis on psychological burdens and

motivators. Regardless, this is consistent with other literature on uptake and motivation, which has only found psychological barriers and facilitators of persistence with interventions [42,44,45].

Figure 1. Conceptualization of the relationship between the barriers and motivating factors that influence persistence in the Cardiovascular Risk E-couch Depression Outcome (CREDO) research project.



Limitations

First and most important, these findings, like all qualitative research, are unique to this study population and the analysis of this research team, and the analysis of these data was influenced by the perspective of the researchers. If another team, with a different theoretical position, were to analyze the transcripts, they would likely find different results. The primary

author has worked in the area of medication adherence and also works as a clinical and health psychologist, while the secondary author is a psychiatrist with a background in epidemiology. These positions have resulted in the development of views of what influences adherence behavior in a clinical setting. This, however, may be different from the influences of adherence in a research or online forum. Given this, we attempted to be as

objective as possible, but the analysis is likely to have been couched within these experiences.

Second, the participants interviewed were people who volunteered to participate in a qualitative interview, from a group that volunteered to participate in an online study, from a group that volunteered to participate in a longitudinal study of health. Therefore, this group was likely to be highly motivated to participate in research and may have been different from those who did not wish to share their views. This played out in the reported emphasis of research as an important motivator by all interview participants.

Due to trial logistic constraints, participants were not contacted until approximately 9 months after they commenced the intervention, or 6 months after they had completed it. This meant that several of the participants may have struggled to recall their intrinsic motivations for continuing to engage in the trial, and their recall was likely influenced by their present situation in order to compensate for the information that they had forgotten [54]. While some of the themes that have emerged appear to be values based, and therefore quite static, it is likely the state variables such as time factors, mood, and frustration may not be as accurately reflected in these findings.

Finally, we were not able to recruit anyone who had not completed the trial or withdrawn, preventing us from exploring the barriers and motivators in people who do not persist with the intervention. Of those who did not persist and who did provide additional contact with the researchers, their reasons for dropping out of the trial tended to comprise a lack of time to complete the program at present, and they were therefore unwilling to participate in a further interview. Future research would benefit from interviewing participants who dropped out of online interventions to determine their reasons for doing so.

Clinical Implications

In terms of maintaining participants in online programs, it may be useful to consider ways to maximize intrinsic motivations and overcome the barriers (through extrinsic motivators) that participants mentioned. This needs to occur in order to meet the dynamic state of motivation. Given this, we propose four steps.

First, briefly, initial hope could be enhanced through educating users about the benefits of Internet interventions and testimonies from previous users about the program. The use of a hook (a message or program design to build curiosity) to engage participants who are starting the intervention due to interest rather than hope for improvement may form a prelude to this and engage users early on.

Second, ongoing engagement could be enhanced through using free-text boxes and providing feedback to activities. While researchers have traditionally been cautious about free text due to the need to manage risk, participants appear to want this to enhance their experience of using the program. These aspects need to be considered if online interventions are to be used to treat people in the community without the intensive follow-up and potential obligation of a research trial.

Third, time barriers could be overcome through the use of habit-forming strategies such as having the program scheduled to be completed each week. Such scheduling would need to be balanced to ensure that program flexibility is not lost, but rather that participants could nominate a time to complete the program next week and be assigned a reminder service for this. Users could be educated about the importance of completing the program regularly and prioritizing this, in a manner similar to how they are educated about taking medication correctly. The findings from this study indicate that online interventions benefit significantly from the use of reminder systems to prompt participants who have forgotten to continue with the interventions.

Fourth, motivation could be enhanced through messaging throughout the program about the benefits of completing the program on a regular basis. Participants may benefit from personal reinforcement through having their outcome scores tracked over time and receiving tailored feedback on this. This may be particularly useful for patients who may struggle to perceive their personal improvements over the course of the program. This type of feedback needs to be provided repeatedly to keep participants motivated and to highlight their achievements and progress through the course of the program. Such an approach is similar to building mastery, which has been shown to enhance persistence with educational study [55] and with therapy [56,57]. Messaging that builds on values of completion or supporting others would also be useful. While this is somewhat difficult to generate in a public health or open access setting, the highlighting of constant monitoring and refinement of online interventions based on participant behavior and feedback may enhance program persistence for people who are completing the program for the benefit of the majority, rather than for personal gain.

However, such communication would benefit from being delivered on a variable-ratio schedule to decrease the predictability of messaging and increase interest in the communications. Further knowledge about building engagement and persistence could be drawn from the self-help book literature to further explore this area of research [58]. Finally, behavioral economic manipulation to enhance the perceived relative value of the program may also increase engagement.

This study is unique in that participants were older adults with physical health comorbidities, and little is known about the use of online interventions in such groups. In addition, the intervention consisted of 12 weekly modules, longer than the 9-module intervention studied by Gerhards et al [44] and the 8-module program used by Bendelin et al [45], thereby requiring greater persistence and motivation from the participants to complete the intervention. While the findings of this study are consistent with these earlier studies [44,45], our study proposes that persistence can be enhanced through understanding and addressing four dynamic processes: (1) building initial hope for benefits of the program, (2) enhancing personal engagement, (3) reducing barriers to use, and (4) increasing ongoing motivating factors. Future research would benefit from evaluating manipulation of the processes identified in this study to determine how these influence persistence with e-therapy interventions and the degree to which this improves outcomes.

Acknowledgments

This CREDO research trial is supported by the Cardiovascular Disease and Depression Strategic Research Program (Award Reference No. G08S 4048), funded by the National Heart Foundation of Australia, and beyondblue: the national depression initiative.

The 45 and Up Study is managed by The Sax Institute in collaboration with major partner Cancer Council New South Wales, as well as partners the National Heart Foundation of Australia (NSW Division); NSW Health; beyondblue: the national depression initiative; Ageing, Disability and Home Care, Department of Human Services NSW; and UnitingCare Ageing.

Conflicts of Interest

None declared.

References

1. Wootton BM, Titov N, Dear BF, Spence J, Kemp A. The acceptability of Internet-based treatment and characteristics of an adult sample with obsessive compulsive disorder: an Internet survey. *PLoS One* 2011;6(6):e20548 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0020548](#)] [Medline: [21673987](#)]
2. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. *Psychol Med* 2008 Nov;38(11):1521-1530. [doi: [10.1017/S0033291707002607](#)] [Medline: [18205964](#)]
3. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. *Australas Psychiatry* 2011 Jun;19(3):259-264. [doi: [10.3109/10398562.2011.562295](#)] [Medline: [21682626](#)]
4. Cavanagh K, Shapiro DA, Van Den Berg S, Swain S, Barkham M, Proudfoot J. The acceptability of computer-aided cognitive behavioural therapy: a pragmatic study. *Cogn Behav Ther* 2009 Dec;38(4):235-246. [doi: [10.1080/16506070802561256](#)] [Medline: [19306147](#)]
5. Titov N, Andrews G, Johnston L, Schwencke G, Choi I. Shyness programme: longer term benefits, cost-effectiveness, and acceptability. *Aust N Z J Psychiatry* 2009 Jan;43(1):36-44. [doi: [10.1080/00048670802534424](#)] [Medline: [19085526](#)]
6. 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](#)]
7. Griffiths K, Christensen H. Review of randomised controlled trials of Internet interventions for mental disorders and related conditions. *Clin Psychol* 2006;10(1):16-29. [doi: [10.1080/13284200500378696](#)]
8. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. *J Med Internet Res* 2004 Nov 10;6(4):e40 [[FREE Full text](#)] [doi: [10.2196/jmir.6.4.e40](#)] [Medline: [15631964](#)]
9. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. *Cochrane Database Syst Rev* 2005(4):CD004274. [doi: [10.1002/14651858.CD004274.pub4](#)] [Medline: [16235356](#)]
10. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. *Ann Behav Med* 2009 Aug;38(1):18-27. [doi: [10.1007/s12160-009-9133-4](#)] [Medline: [19802647](#)]
11. Haaga DA. Introduction to the special section on stepped care models in psychotherapy. *J Consult Clin Psychol* 2000 Aug;68(4):547-548. [Medline: [10965628](#)]
12. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost? *Expert Rev Pharmacoecon Outcomes Res* 2007 Jun;7(3):291-297. [doi: [10.1586/14737167.7.3.291](#)] [Medline: [20528315](#)]
13. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Nuro KF, et al. Computer-assisted delivery of cognitive-behavioral therapy for addiction: a randomized trial of CBT4CBT. *Am J Psychiatry* 2008 Jul;165(7):881-888. [doi: [10.1176/appi.ajp.2008.07111835](#)] [Medline: [18450927](#)]
14. Titov N, Andrews G, Schwencke G, Drobny J, Einstein D. Shyness 1: distance treatment of social phobia over the Internet. *Aust N Z J Psychiatry* 2008 Jul;42(7):585-594. [doi: [10.1080/00048670802119762](#)] [Medline: [18612862](#)]
15. Titov N, Andrews G, Schwencke G. Shyness 2: treating social phobia online: replication and extension. *Aust N Z J Psychiatry* 2008 Jul;42(7):595-605. [doi: [10.1080/00048670802119820](#)] [Medline: [18612863](#)]
16. Taylor CB, Jobson KO, Winzelberg A, Abascal L. The use of the Internet to provide evidence-based integrated treatment programs for mental health. *Psychiatr Ann* 2002 Nov;32(11):671-677 [[FREE Full text](#)]
17. Houston TK, Cooper LA, Vu HT, Kahn J, Toser J, Ford DE. Screening the public for depression through the Internet. *Psychiatr Serv* 2001 Mar;52(3):362-367. [Medline: [11239106](#)]
18. Proudfoot JG. Computer-based treatment for anxiety and depression: is it feasible? Is it effective? *Neurosci Biobehav Rev* 2004 May;28(3):353-363. [doi: [10.1016/j.neubiorev.2004.03.008](#)] [Medline: [15225977](#)]

19. Kenwright M, Marks I, Marks IM, Gega L, Mataix-Cols D. Computer-aided self-help for phobia/panic via internet at home: a pilot study. *Br J Psychiatry* 2004 May;184:448-449 [[FREE Full text](#)] [Medline: [15123511](#)]
20. Leach LS, Christensen H, Griffiths KM, Jorm AF, Mackinnon AJ. Websites as a mode of delivering mental health information: perceptions from the Australian public. *Soc Psychiatry Psychiatr Epidemiol* 2007 Feb;42(2):167-172. [doi: [10.1007/s00127-006-0138-z](#)] [Medline: [17160594](#)]
21. Haemmerli K, Znoj H, Berger T. Internet-based support for infertile patients: a randomized controlled study. *J Behav Med* 2010 Apr;33(2):135-146. [doi: [10.1007/s10865-009-9243-2](#)] [Medline: [20039196](#)]
22. Andersson G, Carlbring P, Grimlund A. Predicting treatment outcome in internet versus face to face treatment of panic disorder. *Comput Hum Behav* 2008;24(5):1790-1801. [doi: [10.1016/j.chb.2008.02.003](#)]
23. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. *J Med Internet Res* 2006;8(2):e10 [[FREE Full text](#)] [doi: [10.2196/jmir.8.2.e10](#)] [Medline: [16867965](#)]
24. Cavanagh K, Shapiro DA. Computer treatment for common mental health problems. *J Clin Psychol* 2004 Mar;60(3):239-251. [doi: [10.1002/jclp.10261](#)] [Medline: [14981789](#)]
25. Barak A, Hen L, Boniel-Nissim M, Shapira M. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. *Technol Hum Serv* 2008;26(2):109-160. [doi: [10.1080/15228830802094429](#)]
26. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009 Dec;38(4):196-205. [doi: [10.1080/16506070903318960](#)] [Medline: [20183695](#)]
27. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. *Clin Psychol Rev* 2009 Jul;29(5):410-420. [doi: [10.1016/j.cpr.2009.04.001](#)] [Medline: [19403215](#)]
28. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009;11(2):e13 [[FREE Full text](#)] [doi: [10.2196/jmir.1194](#)] [Medline: [19403466](#)]
29. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002 Mar 2;359(9308):781-785. [doi: [10.1016/S0140-6736\(02\)07882-0](#)] [Medline: [11888606](#)]
30. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 2005 Feb;34(1):79-87 [[FREE Full text](#)] [doi: [10.1093/ije/dyh300](#)] [Medline: [15561753](#)]
31. 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](#)]
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. Batterham PJ, Neil AL, Bennett K, Griffiths KM, Christensen H. Predictors of adherence among community users of a cognitive behavior therapy website. *Patient Prefer Adherence* 2008;2:97-105 [[FREE Full text](#)] [Medline: [19920949](#)]
34. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *J Med Internet Res* 2004 Dec 22;6(4):e46 [[FREE Full text](#)] [doi: [10.2196/jmir.6.4.e46](#)] [Medline: [15631970](#)]
35. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. *J Med Internet Res* 2005;7(1):e7 [[FREE Full text](#)] [doi: [10.2196/jmir.7.1.e7](#)] [Medline: [15829479](#)]
36. Titov N, Andrews G, Choi I, Schwencke G, Mahoney A. Shyness 3: randomized controlled trial of guided versus unguided Internet-based CBT for social phobia. *Aust N Z J Psychiatry* 2008 Dec;42(12):1030-1040. [doi: [10.1080/00048670802512107](#)] [Medline: [19016091](#)]
37. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrlander C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. *J Med Internet Res* 2010;12(1):e3 [[FREE Full text](#)] [doi: [10.2196/jmir.1361](#)] [Medline: [20147006](#)]
38. Jeffery RW, Danaher BG, Killen J, Farquhar JW, Kinnier R. Self-administered programs for health behavior change: smoking cessation and weight reduction by mail. *Addict Behav* 1982;7(1):57-63. [Medline: [7080885](#)]
39. Celio AA, Winzelberg AJ, Dev P, Taylor CB. Improving compliance in on-line, structured self-help programs: evaluation of an eating disorder prevention program. *J Psychiatr Pract* 2002 Jan;8(1):14-20. [Medline: [15985850](#)]
40. Ritterband LM, Ardan K, Thorndike FP, Magee JC, Saylor DK, Cox DJ, et al. Real world use of an Internet intervention for pediatric encopresis. *J Med Internet Res* 2008;10(2):e16 [[FREE Full text](#)] [doi: [10.2196/jmir.1081](#)] [Medline: [18653440](#)]
41. Anhøj J, Jensen AH. Using the internet for life style changes in diet and physical activity: a feasibility study. *J Med Internet Res* 2004 Sep 8;6(3):e28 [[FREE Full text](#)] [doi: [10.2196/jmir.6.3.e28](#)] [Medline: [15471754](#)]
42. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. *Psychol Med* 2009 May;39(5):705-712. [doi: [10.1017/S0033291708004224](#)] [Medline: [18812006](#)]
43. Andersson G, Carlbring P, Berger T, Almlöv J, Cuijpers P. What makes Internet therapy work? Towards a science of cognitive, emotional and behavioural change. *Cogn Behav Ther* 2009;38(suppl 1):55-60. [doi: [10.1080/16506070902916400](#)]

44. Gerhards SA, Abma TA, Arntz A, de Graaf LE, Evers SM, Huibers MJ, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. *J Affect Disord* 2011 Mar;129(1-3):117-125. [doi: [10.1016/j.jad.2010.09.012](https://doi.org/10.1016/j.jad.2010.09.012)] [Medline: [20889214](#)]
45. Bendelin N, Hesser H, Dahl J, Carlbring P, Nelson KZ, Andersson G. Experiences of guided Internet-based cognitive-behavioural treatment for depression: a qualitative study. *BMC Psychiatry* 2011;11:107 [FREE Full text] [doi: [10.1186/1471-244X-11-107](https://doi.org/10.1186/1471-244X-11-107)] [Medline: [21718523](#)]
46. Cockayne NL, Glozier N, Naismith SL, Christensen H, Neal B, Hickie IB. Internet-based treatment for older adults with depression and co-morbid cardiovascular disease: protocol for a randomised, double-blind, placebo controlled trial. *BMC Psychiatry* 2011;11:10 [FREE Full text] [doi: [10.1186/1471-244X-11-10](https://doi.org/10.1186/1471-244X-11-10)] [Medline: [21232162](#)]
47. Bloor M, Wood F. *Keywords in Qualitative Methods: A Vocabulary of Research Concepts*. London: Sage Publications; 2006.
48. 45 and Up Study Collaborators, Banks E, Redman S, Jorm L, Armstrong B, Bauman A, et al. Cohort profile: the 45 and up study. *Int J Epidemiol* 2008 Oct;37(5):941-947 [FREE Full text] [doi: [10.1093/ije/dym184](https://doi.org/10.1093/ije/dym184)] [Medline: [17881411](#)]
49. Charmaz KC. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: Sage Publications; 2006.
50. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *J Psychosom Res* 1999 Dec;47(6):555-567. [Medline: [10661603](#)]
51. Waber RL, Shiv B, Carmon Z, Ariely D. Commercial features of placebo and therapeutic efficacy. *JAMA* 2008 Mar 5;299(9):1016-1017. [doi: [10.1001/jama.299.9.1016](https://doi.org/10.1001/jama.299.9.1016)] [Medline: [18319411](#)]
52. Madsen SM, Mirza MR, Holm S, Hilsted KL, Kampmann K, Riis P. Attitudes towards clinical research amongst participants and nonparticipants. *J Intern Med* 2002 Feb;251(2):156-168. [Medline: [11905591](#)]
53. Beattie A, Shaw A, Kaur S, Kessler D. Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. *Health Expect* 2009 Mar;12(1):45-59. [doi: [10.1111/j.1369-7625.2008.00531.x](https://doi.org/10.1111/j.1369-7625.2008.00531.x)] [Medline: [19250152](#)]
54. Harvey AG, Watkins E, Mansell W, Shafran R. *Cognitive Behavioural Processes Across Psychological Disorders: A Transdiagnostic Approach to Research and Treatment*. Oxford: Oxford University Press; 2004.
55. Harackiewicz JM, Barron KE, Tauer JM, Carter SM, Elliot AJ. Short-term and long-term consequences of achievement goals: predicting interest and performance over time. *J Educ Psychol* 2000;92(2):316-330. [doi: [10.1037/0022-0663.92.2.316](https://doi.org/10.1037/0022-0663.92.2.316)]
56. Longo DA, Lent RW, Brown SD. Social cognitive variables in the prediction of client motivation and attrition. *J Couns Psychol* 1992;39(4):447-452. [doi: [10.1037/0022-0167.39.4.447](https://doi.org/10.1037/0022-0167.39.4.447)]
57. Sharf J. *Psychotherapy Dropout: A Meta-analytic Review of Premature Termination* [thesis]. Garden City, NY: Adelphi University, Institute of Advanced Psychological Studies; 2007.
58. Richardson R, Richards DA, Barkham M. Self-help books for people with depression: the role of the therapeutic relationship. *Behav Cogn Psychother* 2010 Jan;38(1):67-81. [doi: [10.1017/S1352465809990452](https://doi.org/10.1017/S1352465809990452)] [Medline: [19995466](#)]

Abbreviations

CREDO: Cardiovascular Risk E-couch Depression Outcome

Edited by G Eysenbach; submitted 05.03.12; peer-reviewed by G Andersson, S Langrial; comments to author 26.03.12; revised version received 04.05.12; accepted 11.06.12; published 22.06.12.

Please cite as:

Donkin L, Glozier N

Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers
J Med Internet Res 2012;14(3):e91

URL: <http://www.jmir.org/2012/3/e91/>

doi: [10.2196/jmir.2100](https://doi.org/10.2196/jmir.2100)

PMID: [22743581](https://pubmed.ncbi.nlm.nih.gov/22743581/)

©Liesje Donkin, Nick Glozier. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 22.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of an Online Group Course for Depression in Adolescents and Young Adults: A Randomized Trial

Rianne van der Zanden¹, MSc; Jeannet Kramer², PhD; Rob Gerrits³, MSc; Pim Cuijpers⁴, PhD

¹Centre of Mental Health of Youth and Adolescents, Trimbos Institute, Utrecht, Netherlands

²Innovation Centre of Mental Health and Technology, Trimbos Institute, Utrecht, Netherlands

³Dimence Institute of Mental Health, Deventer, Netherlands

⁴Department of Clinical Psychology, EMGO Institute for Health and Care Research, Vrije Universiteit, Amsterdam, Netherlands

Corresponding Author:

Rianne van der Zanden, MSc

Centre of Mental Health of Youth and Adolescents

Trimbos Institute

Da Costakade 45

Utrecht, 3500 AS Utrecht

Netherlands

Phone: 31 030 2971100

Fax: 31 030 297111

Email: rzanden@trimbos.nl

Abstract

Background: Depression is a serious mental health problem, whose first onset is usually in adolescence. Online treatment may offer a solution for the current undertreatment of depression in youth. For adults with depressive symptoms, the effectiveness of Internet-based cognitive behavioral therapy has been demonstrated. This study is one of the first randomized controlled trials to investigate the effectiveness online depression treatment for young people with depressive complaints and the first to focus on an online group course.

Objective: To evaluate and discuss the effectiveness of a guided Web-based group course called Grip op Je Dip (Master Your Mood [MYM]), designed for young people aged 16 to 25 years with depressive symptoms, in comparison with a wait-listed control group.

Methods: We randomly assigned 244 young people with depressive symptoms to the online MYM course or to a waiting-list control condition. The primary outcome measure was treatment outcome after 3 months on the Center for Epidemiologic Studies Depression Scale. Secondary outcomes were anxiety (measured by the Hospital Anxiety and Depression Scale) and mastery (Mastery Scale). We studied the maintenance of effects in the MYM group 6 months after baseline. Missing data were imputed.

Results: The MYM group ($n = 121$) showed significantly greater improvement in depressive symptoms at 3 months than the control group ($n = 123$) ($t_{187} = 6.62$, $P < .001$), with a large between-group effect size of $d = 0.94$ (95% confidence interval [CI] 0.64–1.23). The MYM group also showed greater improvement in anxiety ($t_{187} = 3.80$, $P < .001$, $d = 0.49$, 95% CI 0.24–0.75) and mastery ($t_{187} = 3.36$, $P = .001$, $d = 0.44$, 95% CI 0.19–0.70). At 12 weeks, 56% (68/121) of the participants in the MYM group and 20% (24/123) in the control group showed reliable and clinically significant change. This between-group difference was significant ($\chi^2_1 = 35.0$, $P < .001$) and yielded a number needed to treat of 2.7. Improvements in the MYM group were maintained at 6 months. A limitation is the infeasibility of comparing the 6-month outcomes of the MYM and control groups, as the controls had access to MYM after 3 months.

Conclusions: The online group course MYM was effective in reducing depressive symptoms and anxiety and in increasing mastery in young people. These effects persisted in the MYM group at 6 months.

Trial Registration: Netherlands Trial Register: NTR1694; <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=1694> (Archived by WebCite at <http://www.webcitation.org/683SBoeGV>)

(*J Med Internet Res* 2012;14(3):e86) doi:[10.2196/jmir.2033](https://doi.org/10.2196/jmir.2033)

KEYWORDS

eHealth, health promotion; depressive symptoms; anxiety; adolescents; Internet; randomized controlled trial

Introduction

Depression Among Adolescents

Depression is a major health problem. Worldwide it is the fourth-ranked disorder in terms of disease burden, and it is expected to carry the highest disease burden in high-income countries by 2030 [1]. The 12-month prevalence of depression is now 5.5% in high-income countries [2]. It is a common condition in adolescents and young adults. One study found a 12-month prevalence of 6.7% among 18- to 25-year-olds [3]. Adolescent depression is associated with serious problems, including poor school performance [4], school absenteeism and dropout [5], problematic relations with parents and peers [6], excessive tobacco and alcohol use [7], and suicidal behavior [8].

Subclinical depression is also common, with an estimated prevalence of 17% to 21% among Dutch adolescents [9,10]. It involves having some depressive symptoms that together do not meet the full *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition criteria for major depression [11]. Subclinical depression is a risk factor for the development of major depression within a year [12,13]. The psychosocial functional impairment of people with subclinical depression is comparable with that in those who have a diagnosis of major depression [14,15]. Beyond the personal suffering involved, both major and subthreshold depression impose significant economic burdens in terms of health care costs and production losses in paid and unpaid work [16,17].

Given the high prevalence rates, serious outcomes, and economic burden, the World Health Organization calls for the development of preventive interventions to reduce the burden of this disorder [18,19]. The first onset is usually in adolescence [20], and it is wise to intervene at an early stage. Yet young people experience many barriers to seeking professional help. They tend to deny or underestimate the problems, fear stigmatization, and question the benefits of help [21]. If they do seek help, they often encounter waiting lists [22,23].

Internet-Based Interventions for Depression

By offering a solution to the stigmatization problem, Internet-based approaches could help in reaching target groups who might otherwise remain untreated. The Internet provides anonymity and the opportunity to take part in an intervention in the privacy of the home. Another strong benefit is that Internet-based approaches enable a reduction in contact hours between professionals and clients, which could help tackle the problem of waiting lists, shortages of therapists, and rising health care costs. Web-based interventions with professional support have been found effective in treating depression in adults, with results comparable with traditional psychological approaches [24-26].

The present study focused on one specific type of Web-based intervention: a professionally facilitated, cognitive behavioral therapy (CBT) group course designed for young people with

symptoms of depression. The perceived advantages of online group sessions as compared with individual approaches are social support and mutual recognition by group members (though they remain anonymous to one another) and the reduction of professional contact hours per participant in comparison with individual treatment [27]. Three earlier studies on online group courses in mental health care [28-30] showed that it is possible, both technically and substantively, to conduct Internet-based group courses for adolescents and adults with mental health problems via the Internet. Two of these three studies showed significant positive outcomes in pre-post measurements, though also reporting relatively high attrition rates—a common problem for Internet interventions [28]. The pilot study by Gerrits et al [29] among 140 adolescents with depressive complaints who received the online CBT group course Master Your Mood (MYM) showed a significant decline in depressive symptoms. And the pilot study by Van der Zanden et al [30] of 48 parents with mental illness who received parenting support in an online group course showed significant improvement in parenting skills and parental sense of competence.

Lack of Outcome Studies for Youth

Despite the clear benefits of Web-based interventions for depression, there is a lack of outcome studies specifically focusing on adolescents and young adults [31,32]. Only two randomized controlled trials have been conducted on prevention programs for depression and anxiety [33,34]. Calear and colleagues [33] found effects of a universal prevention program for male and female participants for anxiety but found effects for depression only in male participants. Van Voorhees et al [34] showed that a Web-based behavior change program in primary care, in combination with either motivational interviewing or brief advice, was associated with declines in depressed mood and levels of depressive symptoms.

Objective

We evaluate and discuss the effectiveness of a Web-based group course called Grip op Je Dip (Master Your Mood), designed for young people aged 16 to 25 years with depressive symptoms, in comparison with a wait-listed control group. The primary outcome measure was depression, and secondary outcomes were anxiety and sense of control. Based on the results of the pilot study [29] we expected better outcomes for the course group.

Methods

Study Design

We conducted a randomized controlled trial with two parallel groups to examine the effectiveness of the MYM course, comparing the intervention group with a waiting-list control group. Ethical approval was granted by an independent medical ethics committee (CCMO no. NL18984.097.07). The trial is registered (NTR1694), and the study protocol has been published [35].

Participating Mental Health Agencies

A total of 14 mental health care agencies participated in the project, all working with online course participants from all over the Netherlands. The courses were supervised by professional mental health promotion workers, trained in administering the Mini-International Neuropsychiatric Interview (MINI-Plus) [36,37] and in conducting the course.

Study Population

The inclusion criteria were the following: age 16 to 25 years, informed consent (including parental consent for those less than 18 years of age), and a Center for Epidemiologic Studies Depression Scale (CES-D) score between 10 and 45. Applicants were excluded on indications of suicidal ideation with intent and plan, as assessed with the MINI-Plus.

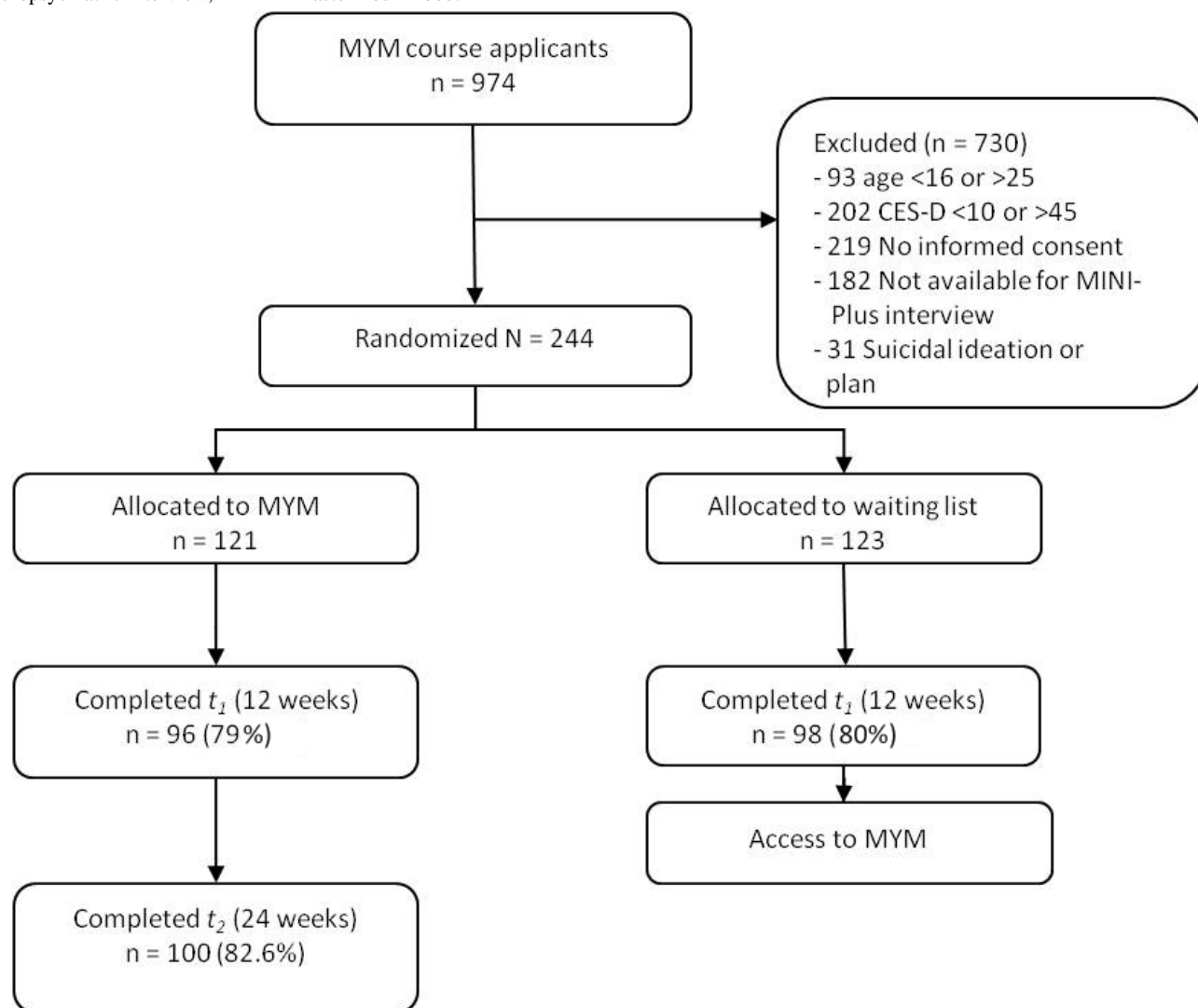
Recruitment Procedure

Participants were recruited from the general population by means of promotional materials in general practitioners' offices and educational institutions. Banners and links were also placed on mental health-related websites and on websites popular with young people. There were no explicit restrictions on country of origin, but the course language was Dutch and the recruitment took place in Dutch. The minimum requirements for Internet access were a stable Internet connection and a recent browser (minimum: Internet Explorer 6.0). Participants also had to be able to read, write, and chat in Dutch on at least the primary school level. Those interested were referred to the MYM website

[38] to complete an online preliminary screening questionnaire and apply for the course. Those with CES-D scores between 10 and 45 then received additional information about the study, an informed consent form (including a parental consent form for 16- and 17-year-olds), and a baseline questionnaire. For course applicants scoring 25 to 45 on the CES-D, a mandatory online chat session followed, in which suicidal ideation was assessed with the MINI-Plus interview. Those for whom suicidal ideation and plan were determined were excluded from the study and advised to see their general practitioner.

Eligible applicants were randomly assigned to the intervention group (MYM) or the control group (wait-listed for 14 weeks). Random allocation was automated by a computer program with no interference by course facilitators or researchers. A blocked randomization scheme was used with blocks of two, stratified by depressive symptoms (CES-D scores of 10–24 vs 25–45) and age (younger vs older than 18 years). The outcome of the randomization was generated and made available at the moment the course facilitator indicated that the applicant was eligible for the course. Applicants were informed of their allocation by email and received a tailored referral if they were declined. Participants then received a personal email from their facilitator to inform them of the specific times and dates of the course, the homework assignment for the first session, and a username and password for the chat room. During the trial, participants in both conditions were allowed to seek additional help if they wished. [Figure 1](#) shows a flow chart of respondent selection.

Figure 1. Flow chart of respondent selection. CES-D = Center for Epidemiologic Studies Depression Scale; MINI-Plus: Mini-International Neuropsychiatric Interview; MYM = Master Your Mood.



Conditions: The Intervention

The online MYM course is based on the face-to-face intervention of the same name, which was developed by the Trimbos Institute, the Netherlands Institute of Mental Health and Addiction. That intervention was derived from the Dutch version [39] of the Coping with Depression course [40]. The face-to-face course was adapted to the Internet in a collaborative project involving the Trimbos Institute and three mental health agencies [29,41].

The online MYM group course is a structured form of CBT for depression. At the core of MYM is the cognitive restructuring of thinking patterns. Course participants are encouraged to detect their own unproductive, unrealistic thoughts, and they are then taught to transform these into realistic, helpful thoughts. Performance of pleasant daily activities is also encouraged, and

a mood measure is filled in daily to help understand the connection between pleasant activities and mood level.

The course we evaluated took place at fixed times in a secured chat room, which participants entered with their username and password. Anonymity within the group was ensured by a self-chosen nickname. Text messages reminders were sent to participants' mobile phones one-half hour before each session. The course comprised six sessions of 90 minutes each, each at a set time every week, and home exercises. The sessions were structured around six themes (see Table 1). During the sessions, course material was introduced by the facilitators and displayed in the chat room using text and images. Participants could respond, share experiences, and ask questions. Emoticons could be used to express feelings. Participants and professionals could read through the session transcripts afterward. The course was guided by one or two trained professionals, depending on group size (6 participants was the maximum).

Table 1. Outline of the online course Master Your Mood.

Session	Actions
1 Your mood	Learning to understand the relationship between feelings, thoughts, and actions.
2 Your mood and being active	Becoming aware of the influence that activities have on your mood and starting to be more active.
3 Tracing negative thoughts	Becoming aware of negative thoughts and starting to understand what they do to your mood and self-esteem.
4 New ways of thinking	Challenging the negative thoughts and changing them into more positive (but realistic) ones.
5 More action with positive thinking	Combining the new way of thinking with being more active. Becoming more aware of positive things about yourself and others.
6 The future	Making a plan to prevent relapse into depressive moods in the future. Learning the personal signals of a coming depression and knowing how to address that threat. Making a personal plan for the future, to include wishes about schooling, jobs, and relationships and ways to achieve those wishes.

Conditions: The Waiting List

The wait-listed group did not receive an intervention. They were told by email that they would be invited to participate after the waiting period of 3 months.

Assessments

Assessments took place before randomization (baseline, $t0$), 12 weeks later ($t1$), and 12 weeks after that ($t2$). Participants received automated emails with invitations to complete the online questionnaires. Subsequent reminders were sent 5 and 10 days after the first email invitation, if necessary. To stimulate response, we offered the participants €10 compensation for completing the $t1$ questionnaire and €10 for the $t2$ questionnaire, plus an additional €5 bonus (totaling €25) if they completed both.

Measures

Primary Outcome Measure: Depressive Symptoms

The CES-D [42,43] measures the frequency of 20 depressive symptoms over the past week on a 4-point Likert scale. The total score may range from 0 to 60, with higher scores indicating higher levels of depression. Computerized and paper-and-pencil versions of the CES-D correlate at a very high level [44]. The Web-based version of the CES-D has been shown to be a reliable and valid screening instrument in a Dutch adolescent population, with a Cronbach alpha of 0.93 [45]. In our study, the Cronbach alpha was 0.91.

Secondary Outcome Measures

Anxiety

We used the Anxiety subscale of the Hospital Anxiety and Depression Scale (HADS) to assess anxiety symptoms [46]; the Dutch version of the HADS has been validated [47]. The Cronbach alpha was 0.84 for adults (18+ years) in the general population. The Anxiety subscale consists of 7 items measuring symptoms of anxiety on a 4-point Likert scale, with a score range of 0 to 21 and a higher rating indicating a higher state of anxiety. In our study the Cronbach alpha of this subscale was 0.74.

Sense of Control

We used the Dutch version of the 5-item Mastery Scale [48] to assess perceived control. The concept of mastery refers to beliefs about one's own ability to control one's environment. Responses

are rated on a 5-point Likert scale, with a total score range of 5 to 25; a higher score indicates a greater sense of mastery. The Mastery Scale has good psychometric properties [48]. In our study, the Cronbach alpha was 0.77.

Additional Measures

At baseline, we assessed demographic information (sex, age, educational level, and living situation), previous or present professional help for psychological problems, and experience with Web chatting. At 3 months ($t1$) and 6 months ($t2$), we inquired about subsequent use of professional help and antidepressant medication.

Analyses

The trial was originally powered to detect a clinical effect of $d = 0.32$ or larger in a 1-sided test with a power of 80% ($1 - \beta$) [29]. Hypotheses were directional, with better outcomes expected for the MYM group. A total of 242 participants were needed for the study, $n = 121$ per condition (Stata 11.1 syntax: `sampsi`; StataCorp LP, College Station, TX, USA).

We used t tests, chi-square tests, and logistic regression ($P < .05$) to determine whether the randomization had resulted in two comparable groups at baseline and whether any differential loss to follow-up occurred.

The expectation-maximization method was used to impute data missing at $t1$ and $t2$. It imputes values by maximum-likelihood estimation using the observed data in an iterative process [49]. These analyses were based on the intention-to-treat principle, including data from all participants, whether or not they received the intervention.

To test course effectiveness, we calculated difference scores between $t0$ and $t1$ for all outcome variables (positive scores meaning improvement) and compared them across groups in linear regression models, controlling for data clustering (some participants attended the same course sessions). Clustering would violate the assumption of independence of observations and might affect standard errors and P values. To adjust for clustering, we obtained robust standard errors and P values using the first-order Taylor series linearization method, as implemented in Stata. At $t2$ (24 weeks), we studied effect maintenance in the MYM group only; we made no between-group comparisons, as the control group had since been given access to the course.

The sizes of intervention effects were estimated using Cohen d [50]. Effect sizes were first calculated for each condition separately, $(t_0 - t_1)/SD_{t_0}$, and then the differential effect size was calculated by subtracting the control group effect from that of the MYM group. A difference of $d = 0.5$ would indicate that the experimental group mean was half a standard deviation greater than the control group mean. For Cohen d , an effect size of 0.2 to 0.3 may be regarded as a small effect, around 0.5 as a medium effect, and 0.8 to infinity as a large effect.

The proportion of participants showing reliable and clinically significant improvement [51] was determined in terms of an improvement of 5 points on the CES-D in combination with a score lower than 22 on the CES-D (cut-off based on Cuijpers et al [45]). Subsequently, the number needed to treat was calculated as 1/risk difference. The number needed to treat indicates here how many young people with depressive symptoms would need to take the MYM course in order to generate a health gain in 1 person over 12 weeks.

We determined effect maintenance in the MYM group on the basis of the results at 24 weeks (t_2). Paired t tests were used to identify significant changes in outcomes between baseline, 12 weeks (t_1), and 24 weeks in the intention-to-treat sample.

Results

Participants

Recruitment took place from May 20, 2008 to March 6, 2010. Figure 1 shows the flow of participants through the trial. Of the 974 people who applied for the online MYM course, 244 (25.1%) were included in the study. Reasons for noninclusion were lack of informed consent (219/730, 30.0%), CES-D depression score outside the 10–45 range (202/730, 27.7%), no-show at the MINI-Plus interview (182/730, 24.9%), and age outside the 16–25 range (93/730, 13%). Additional exclusions were made for suicidal ideation (31/730, 4%) and other reasons (3/730, 0%). The 244 selected participants were randomly assigned to one of the two conditions: the online MYM course ($n = 121$) and the waiting-list control group ($n = 123$). Table 2 shows baseline demographic and psychosocial characteristics of the 244 participants. Professional help was provided in most cases by a psychologist or other professional from a mental health agency. No significant differences were found on any of these variables between the experimental group and control group ($P < .05$), indicating that the randomization was successful.

Table 2. Baseline characteristics of the 244 participants.

Characteristic	Experimental group (n = 121)	Control group (n = 123)	All (N = 244)	Statistics
Female gender, n (%)	101 (83.5%)	105 (85.4%)	206 (84.4%)	$\chi^2_1 = 0.2, P = .68$
Age (years), mean (SD)	20.8 (2.2)	21.0 (2.3)	20.9 (2.2)	$t_{242} = 0.6, P = .53$
Age range (years), n (%)				$\chi^2_2 = 0.2, P = .92$
16–17	5 (4%)	4 (3%)	9 (4%)	
18–21	66 (55%)	69 (56%)	135 (55.3%)	
22–25	50 (41%)	50 (41%)	100 (41.0%)	
Education level, n (%)^a				$\chi^2_2 = 0.6, P = .73$
Low	10 (8%)	10 (8%)	20 (8%)	
Middle	50 (41%)	45 (37%)	95 (39%)	
High	61 (50%)	68 (55%)	129 (52.8%)	
Daily activities, n (%)				$\chi^2_2 = 1.9, P = .38$
Study	83 (69%)	85 (69%)	168 (69%)	
Paid employment	32 (26%)	27 (22%)	59 (24%)	
Other	6 (5%)	11 (9%)	17 (7%)	
Living situation, n (%)				$\chi^2_3 = 3.5, P = .32$
With parents	56 (46%)	59 (48%)	115 (47.1%)	
With partner	13 (11%)	18 (15%)	31 (13%)	
Alone	26 (22%)	16 (13%)	42 (17%)	
With others	26 (22%)	30 (24%)	56 (23%)	
Experience in Web chat, n (%)	63 (52%)	64 (52%)	127 (52.0%)	$\chi^2_1 = 0.0, P = .99$
Professional help, n (%)				
Prior	75 (62%)	76 (62%)	151 (61.9%)	$\chi^2_1 = 0.0, P = .98$
Current at baseline	36 (30%)	39 (32%)	77 (32%)	$\chi^2_1 = 0.1, P = .74$
Test scores, mean (SD)				
CES-D depression score ^b	32.5 (8.4)	32.3 (8.2)	32.3 (8.3)	$t_{242} = 0.28, P = .77$
HADS Anxiety ^c	11.2 (3.6)	11.8 (3.7)	11.5 (3.6)	$t_{242} = 1.27, P = .21$
Mastery ^d	12.8 (3.4)	12.8 (3.6)	12.8 (3.5)	$t_{242} = 0.17, P = .86$

^a Combination of highest completed or present education: low = primary or lower secondary school or less; middle = intermediate vocational school or secondary school; high = professional school or university.

^b Dutch version of the Center for Epidemiologic Studies Depression Scale [43].

^c Dutch version of the Hospital Anxiety and Depression Scale [47].

^d 5-item Mastery Scale [48].

Attrition

A total of 21% (50/244) of the sample did not complete the assessment at the end of 12 weeks (*t1*). We do not know the reasons for noncompletion of questionnaires. There were no significant differences between groups in completing *t1*. Nor were there significant differences between participants who did and who did not complete the *t1* assessment ($P < .10$). This indicates that loss to follow-up was random. The assessment at

24 weeks (*t2*) was used to study the maintenance of effects in the MYM group and was not completed by 17% (21/121) of that group. There were several significant differences at baseline between the MYM participants who completed *t2* and those who did not. Those who did not return that questionnaire were more likely to be male ($\chi^2_1 = 5.2, P = .02$), to have lower education levels ($\chi^2_2 = 5.3, P = .07$), and to have shown higher

anxiety ($t_{119} = 2.5$, $P = .01$) and sense of mastery ($t_{119} = -1.77$, $P = .08$) at a previous assessment.

Effects of the Intervention

Table 3 shows outcomes in the intention-to-treat sample for the primary (CES-D) and secondary (HADS Anxiety and Mastery) measures as produced by estimation-maximization imputation. The results of the regression analyses with adjustment for clustering were nearly identical to those without adjustment for clustering, indicating the absence of a cluster effect. We therefore present the results of independent-samples t tests. From baseline to 12 weeks, the MYM group showed significantly greater improvement in depressive symptoms, anxiety, and mastery than the control group, with a large effect size for depressive symptoms (CES-D, $d = 0.94$, 95% confidence

interval [CI] 0.64–1.23) and moderate effect sizes for anxiety (HADS Anxiety, $d = 0.49$, 95% CI 0.24–0.75) and mastery ($d = 0.44$, 95% CI 0.19–0.70).

Table 4 shows the outcomes for the subsample completing the $t1$ questionnaires (with no imputation for the MYM or control group). These outcomes scarcely differed from those in Table 3, indicating that imputation had little effect on outcome. Table 5 shows outcomes for the subsample that attended at least one course session compared with the control group, using expectation-maximization imputation for missing values. Compared with the intention-to-treat sample in Table 3, effect sizes were slightly better still, with again a large effect for depressive symptoms (CES-D, $d = 1.13$, 95% CI 0.78–1.47) and moderate effects for anxiety (HADS Anxiety, $d = 0.53$, 95% CI 0.25–0.81) and mastery ($d = 0.51$, 95% CI 0.23–0.79).

Table 3. Effects of Master Your Mood (MYM) course: intention-to-treat analysis of full sample, expectation-maximization imputation.

Instrument	MYM group (n = 121)			Wait-listed controls (n = 123)			Between-group outcomes			
	Baseline ($t0$) mean (SD)	12 weeks ($t1$) mean (SD)	d^a	Baseline ($t0$) mean (SD)	12 weeks ($t1$) mean (SD)	d	Dif d^b	t	df	P value
Depression (CES-D) ^c	32.5 (8.4)	19.3 (9.7)	1.57	32.2 (8.2)	27.0 (8.6)	0.63	0.94	6.62	179.1	<.001
Anxiety (HADS) ^d	11.2 (3.6)	8.0 (3.9)	0.89	11.8 (3.7)	10.4 (3.3)	0.39	0.49	3.80	215.9	<.001
Mastery ^e	12.8 (3.4)	15.9 (4.1)	0.91	12.8 (3.6)	14.5 (3.5)	0.47	0.44	3.36	226.8	.001

^a Individual standardized effect size ($d = t0 - t1 / SDt0$), with positive effect sizes indicating improvement.

^b Effect size differences between MYM group and wait-listed control group.

^c Dutch version of the Center for Epidemiologic Studies Depression Scale [43].

^d Dutch version of the Hospital Anxiety and Depression Scale [47].

^e 5-item Mastery Scale [48].

Table 4. Effects of Master Your Mood (MYM) course: comparison of responding participants only, no imputation.

Instrument	MYM group (n = 96)			Wait-listed controls (n = 98)			Between-group outcomes			
	Baseline ($t0$) mean (SD)	12 weeks ($t1$) mean (SD)	d^a	Baseline ($t0$) mean (SD)	12 weeks ($t1$) mean (SD)	d	Dif d^b	t	df	P value
Depression (CES-D) ^c	32.5 (8.0)	19.1 (10.7)	1.60	31.6 (7.7)	26.7 (9.4)	0.60	0.99	5.69	179.1	<.001
Anxiety (HADS) ^d	10.9 (3.5)	7.8 (4.2)	0.88	11.7 (3.2)	10.2 (3.5)	0.39	0.49	3.10	192	.002
Mastery ^e	12.9 (3.3)	16.1 (4.3)	0.93	12.9 (3.3)	14.6 (3.8)	0.47	0.46	2.82	184.5	.005

^a Individual standardized effect size ($d = t0 - t1 / SDt0$), with positive effect sizes indicating improvement.

^b Effect size differences between MYM group and wait-listed control group.

^c Dutch version of the Center for Epidemiologic Studies Depression Scale [43].

^d Dutch version of the Hospital Anxiety and Depression Scale [47].

^e 5-item Mastery Scale [48].

Table 5. Effects of Master Your Mood (MYM) course: MYM participants attending at least one session compared with wait-listed control group, expectation-maximization imputation.

Instrument	MYM group (n = 96)			Wait-listed controls (n = 123)			Between-group outcomes			
	Baseline (<i>t0</i>)	12 weeks (<i>t1</i>)	<i>d</i> ^a	Baseline (<i>t0</i>)	12 weeks (<i>t1</i>)	<i>d</i>	Dif <i>d</i> ^b	<i>t</i>	<i>df</i>	<i>P</i> value
	mean (SD)	mean (SD)		mean (SD)	mean (SD)					
Depression (CES-D) ^c	33.2 (7.8)	19.6 (10.3)	1.76	32.2 (8.2)	27.0 (8.6)	0.63	1.13	6.15	170.6	<.001
Anxiety (HADS) ^d	11.3 (3.5)	8.6 (4.4)	0.92	11.8 (3.7)	10.4 (3.3)	0.39	0.53	3.53	217	.001
Mastery ^e	12.6 (3.4)	15.8 (4.1)	0.98	12.8 (3.6)	14.5 (3.5)	0.47	0.51	3.36	217	.001

^a Individual standardized effect size ($d = t0 - t1 / SDt1$), with positive effect sizes indicating improvement.

^b Effect size differences between MYM group and wait-listed control group.

^c Dutch version of the Center for Epidemiologic Studies Depression Scale [43].

^d Dutch version of the Hospital Anxiety and Depression Scale [47].

^e 5-item Mastery Scale [48].

Reliable and Clinical Change

At 12 weeks, 56% (68/121) of the participants in the MYM group and 20% (24/123) in the control group showed reliable and clinically significant change (a positive change of 5 points or more on the CES-D in combination with a score below 22). This between-group difference was significant ($\chi^2_1 = 35.0$, $P < .001$) and yielded a number needed to treat of 2.7.

Maintenance of Effects in the MYM Group

Table 6 shows the results of the paired *t* tests for the MYM group (intention-to-treat sample). There was significant improvement in depressive symptoms, anxiety, and mastery from baseline (*t0*) to 12 weeks (*t1*), as well as from baseline to 24 weeks (*t2*). The effect sizes (*d*) were large for all measures. Sense of mastery even improved significantly with a small effect size from 12 weeks (*t1*) to 24 weeks (*t2*). The positive results achieved at 12 weeks were thus maintained at 24 weeks, and mastery continued improving.

Table 6. Effect maintenance in the Master Your Mood group (n = 121).

Instrument	Baseline (<i>t0</i>)	12 weeks (<i>t1</i>)	24 weeks (<i>t2</i>)	<i>t0 – t1</i>		<i>t0 – t2</i>			<i>t1 – t2</i>			
	Mean (SD)	Mean (SD)	Mean (SD)	<i>t a</i>	<i>P</i> value	<i>d b</i>	<i>t</i>	<i>P</i> value	<i>d</i>	<i>t</i>	<i>P</i> value	<i>d</i>
Depression (CES-D) ^c	32.5 (8.4)	19.3 (9.7)	18.3 (10.7)	13.3	<.001	1.57	9.4	<.001	1.69	1.0	.31	0.10
Anxiety (HADS) ^d	11.2 (3.6)	8.0 (3.9)	7.8 (3.8)	12.8	<.001	0.89	10.0	<.001	0.96	0.8	.42	0.06
Mastery ^e	12.8 (3.4)	15.9 (4.1)	16.7 (4.1)	9.2	<.001	0.91	10.4	<.001	1.14	2.8	.006	0.19

^a *df* = 120.

^b Individual standardized effect sizes [$d = (t0 - t1) / SDt0$; $d = (t0 - t2) / SDt0$; $d = (t1 - t2) / SDt1$] with positive effect sizes indicating improvement.

^c Dutch version of the Center for Epidemiologic Studies Depression Scale [43].

^d Dutch version of the Hospital Anxiety and Depression Scale [47].

^e 5-item Mastery Scale [48]. Higher scores stand for higher sense of mastery; differences calculated as $t1 - t0$, $t2 - t0$, $t2 - t1$.

Sessions Attended and Outcome

Not all MYM group participants attended all course sessions: 21% (25/121) did not attend any sessions, 52% (63/121) attended at least four sessions, and 20% (24/121) attended all six. The average number of sessions attended was 3.2 (SD 2.2) with a range from 0 to 6.

Tested at $P < .05$, there were no significant differences in the CES-D mean effect sizes between those attending no sessions ($d = 1.3$) and those attending one or more ($d = 1.6$, $t_{119} = 1.03$, $P = .31$), nor between those attending fewer than 3 ($d = 1.5$) or more than 3 sessions ($d = 1.6$, $t_{110.6} = 0.73$, $P = .47$). Tested at $P < .10$, some differences emerged between participants attending no sessions and those attending at least one. Nonattendees included fewer experienced Web chatters (9/25,

36%) as compared with attendees (54/96, 56%; $\chi^2_1 = 3.3$, $P = .07$), and nonattendees also had lower mean baseline CES-D scores (mean 29.6, SD 10, vs 33.2, SD 7.8; Wald $\chi^2_1 = 3.6$, $P = .06$).

Discussion

Main Results

In this study, the Internet-based CBT group course known as Grip op Je Dip (Master Your Mood) for young people aged 16 to 25 years proved significantly more effective than a waiting-list control group in decreasing depressive symptoms. At 3 months, a large between-group effect size of 0.94 was found. The MYM group also showed a significantly greater reduction in anxiety symptoms (with a medium between-group

effect of 0.49) and improvement in sense of control or mastery (medium effect of 0.44). The proportion of participants showing reliable and clinically significant change was 0.56 in the MYM group versus 0.20 in the control group ($\chi^2_1 = 35.0$, $P < .001$). The reductions in depressive and anxiety symptoms and the increased sense of mastery were maintained in the MYM group at 6-month follow-up.

Comparison With Other Work

Our study is one of the first randomized controlled trials to study online depression treatment for young people [32,33], and it is the first to focus on an online group course [30]. This hampers any solid comparison with prior Internet intervention research, which has chiefly focused on adults and on individual approaches. It is with this limitation in mind that we compare the results of the current study with other work.

The effectiveness of online treatment for depression in adults has already been demonstrated [24-26]. Our study showed the effectiveness of an online intervention for young people with depressive symptoms. The large effect size of 0.94 for the MYM course is comparable with the effects of face-to-face psychotherapy for depression in youth [52-54]. So far, effect size differences for Internet-based adult therapies for depression and anxiety have appeared to be related to the amounts of accompanying therapist support [55]. Since the MYM intervention provided substantial support by the course facilitator in the weekly group sessions, this might partly account for the large effect we obtained. In addition to the significant declines in CES-D scores, the MYM course was also associated with significantly lower scores on the HADS Anxiety subscale. Young people with primary anxiety may therefore also stand to benefit—a possibility supported in a recent trial that showed the effectiveness of an online CBT protocol for adult anxiety disorders and/or depression [56].

We also analyzed clinically significant change in the present study. The proportions of improved and recovered participants (56% in the MYM group vs 20% in the control group) seem in line with those for Web-based interventions for adult depression [57-59].

Dose–effect analysis in our study found no correlation between the number of sessions attended and the intervention outcomes, likewise corresponding to other studies [58-61]. Similarly, a meta-analysis by Weisz et al [62] found treatment duration not to be correlated with outcome. The explanation for this is not clear. In our study we did not ask the participants of the MYM group for their reasons for stopping after one or a few sessions. It could be that those who stopped felt they had recovered enough. This may have leveled out the correlation between sessions attended and outcome. The MYM group also included participants who did not attend any session, and they still displayed intervention effects. An explanation for this might be the difference in study conditions to which the MYM group and the wait-listed group were assigned: although both groups attended no sessions, the MYM participants made an active decision about this, while the wait-listed group did not. Indeed, which specific elements of a treatment are effective and which

exact mechanisms bring on a person's recovery are still at the frontiers of knowledge.

In the current field of Internet-based help, there is an ongoing search for an optimal balance (in terms of outcome and costs) between the relative amounts of therapist support and self-guidance [32]. In comparison with an individual eight-session CBT-based online self-help intervention with professional support [58], the facilitated six-session MYM course entailed less supervision time per participant (160 minutes in the eight-session CBT vs 135 minutes in six-session MYM, based on an average of 4 participants per facilitator), while the between-group effect size in the MYM was larger (0.94 MYM vs 0.72 CBT self-help at 12-week posttest). One possibility is that the group interaction, whereby group members provide some support to one another, might have boosted outcomes over the therapist support.

Limitations

This study has several limitations. One of them is the infeasibility of comparing the 6-month (t_2) outcomes of the MYM and control groups, since the control group had access to the course after t_1 . Another limitation is the passive (waiting-list) control condition. This makes it difficult to conclude that the treatment contained components that were “specifically effective over and above simple compassion, friendliness, attention and belief” [63]. Furthermore, the fact that participants were not blind to their condition, though generally inherent in psychotherapy studies, could have introduced some bias. The exclusive reliance on self-report measures is another limitation; other sources of information were missing. In this study we encountered missing data, though the attrition rate was much lower than in some other studies. High attrition rates are common in studies of Internet-based interventions [64,58]. We found no indication in the analyses that our missing data had affected the results. We dealt with missing values at posttest and follow-up using the expectation-maximization imputation method [49].

Our participants had rather high levels of education relative to the general population, so it is uncertain whether results can be generalized to people with lower education levels. The same can be said for male participants, who were underrepresented in our study in relation to the depression prevalence in the male general population. Adolescents aged 16 and 17 years were also underrepresented, due apparently not to lack of willingness to participate, but to a lack of consent given by parents. Further, we are uncertain whether results can be generalized to people with severe depression, as they were excluded from the study.

Future Research Directions

We have pointed out that there is a lack of randomized controlled trials of Web-based interventions that specifically target adolescents and young adults with depression. From a preventive point of view, research on depression in youth is acutely needed, given the frequent early onset of subclinical and major depression and their far-reaching impacts. Future research should focus on the economic evaluation of Internet-based interventions for youth and on outcome research regarding stepped-care interventions (minimal where possible,

sustainable where necessary). Trials are also needed in which online treatment groups are compared with active online control groups. This could help identify more specific elements of treatment that are effective. People with low socioeconomic status backgrounds are generally underrepresented in study samples but are in particular need due to their higher prevalence of psychological distress [3]. Interventions specially tailored to such groups therefore ought to be developed and studied.

Conclusion

Despite the limitations of the present study, our findings suggest that adolescents and young adults with depressive symptoms

can benefit from an Internet-based CBT group course. In our intervention, with a level of professional support per person roughly equivalent to that in other facilitated Web-based self-help interventions for depression, we found a relatively large effect size. The group aspect of our intervention may have accounted for this. Along with the individual approaches developed so far, online CBT group courses for depression may open additional opportunities to reach people who might otherwise remain untreated, thus achieving long-term cost savings. Future research should focus on economic evaluations of Internet-based interventions for depression in youth.

Acknowledgments

Ethical approval was granted by an independent medical ethics committee (CCMO no. NL18984.097.07).

The study was funded by a research grant from ZonMw (Netherlands Organization for Health Research and Development), grant no. 61300036. No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

Rianne van der Zanden and Rob Gerrits are two of the developers of the online MYM group course, and the Trimbos Institute is a cooperation partner in MYM, but none of these derive financial income from the MYM intervention. There are no competing interests.

Multimedia Appendix 1

CONSORT EHEALTH V1.6 checklist [65].

[PDF File (Adobe PDF File), 586KB - [jmir_v14i3e86_app1.pdf](#)]

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. Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med* 2011;9:90 [FREE Full text] [doi: [10.1186/1741-7015-9-90](#)] [Medline: [21791035](#)]
3. de Graaf R, Ten Have MVD, Van Dorsselaer S. De psychische gezondheid van de Nederlandse bevolking: NEMESIS II: Opzet en de eerste resultaten [The Mental Health of the Dutch Population: Netherlands Mental Health Epidemiological and Incidence study NEMESIS II: Design and First Results]. Utrecht, Netherlands: Trimbos Institute; 2010.
4. Meijer SA, Smit F, Schoemaker CG, Cuijpers P. Gezond verstand: evidence-based preventie van psychische stoornissen [Common Good: Evidence-Based Prevention of Mental Health Disorders]. Bilthoven, Netherlands: RIVM; 2006.
5. Dopheide JA. Recognizing and treating depression in children and adolescents. *Am J Health Syst Pharm* 2006 Feb 1;63(3):233-243. [doi: [10.2146/ajhp050264](#)] [Medline: [16434782](#)]
6. Cicchetti D, Toth SL. The development of depression in children and adolescents. *Am Psychol* 1998 Feb;53(2):221-241. [Medline: [9491749](#)]
7. Glied S, Pine DS. Consequences and correlates of adolescent depression. *Arch Pediatr Adolesc Med* 2002 Oct;156(10):1009-1014. [Medline: [12361447](#)]
8. Fergusson DM, Woodward LJ, Horwood LJ. Risk factors and life processes associated with the onset of suicidal behaviour during adolescence and early adulthood. *Psychol Med* 2000 Jan;30(1):23-39. [Medline: [10722173](#)]
9. Smith F, Bohlmeijer E, Cuijpers P. Wetenschappelijke onderbouwing depressiepreventie. Epidemiologie, aangrijpingspunten, huidige praktijk en nieuwe richtingen [Scientific Foundation of Prevention of Depression: Epidemiology, Pivotal Points, Current Practice and New Directions]. Utrecht, Netherlands: Trimbos Institute; 2003.
10. Ter Bogt T, van Dorsselaer S, Vollebergh W. Psychische gezondheid, risicogedrag en welbevinden van Nederlandse scholieren [Mental Health, Risk Behaviour and Well-being of Dutch Students]. Trimbos Institute: Trimbos Institute; 2003.
11. Cuijpers P, Smit F. Subklinische depressie: een klinisch relevante conditie? [Subclinical depression: a clinically relevant condition?]. *Tijdschr Psychiatr* 2008;50(8):519-528 [FREE Full text] [Medline: [18688776](#)]

12. Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatr Scand* 2004 May;109(5):325-331. [doi: [10.1111/j.1600-0447.2004.00301.x](https://doi.org/10.1111/j.1600-0447.2004.00301.x)] [Medline: [15049768](#)]
13. Georgiades K, Lewinsohn PM, Monroe SM, Seeley JR. Major depressive disorder in adolescence: the role of subthreshold symptoms. *J Am Acad Child Adolesc Psychiatry* 2006 Aug;45(8):936-944. [doi: [10.1097/01.chi.0000223313.25536.47](https://doi.org/10.1097/01.chi.0000223313.25536.47)] [Medline: [16865036](#)]
14. Gotlib IH, Lewinsohn PM, Seeley JR. Symptoms versus a diagnosis of depression: differences in psychosocial functioning. *J Consult Clin Psychol* 1995 Feb;63(1):90-100. [Medline: [7896995](#)]
15. Wagner HR, Burns BJ, Broadhead WE, Yarnall KS, Sigmon A, Gaynes BN. Minor depression in family practice: functional morbidity, co-morbidity, service utilization and outcomes. *Psychol Med* 2000 Nov;30(6):1377-1390. [Medline: [11097078](#)]
16. Cuijpers P, Smit F, Oostenbrink J, de Graaf R, Ten Have M, Beekman A. Economic costs of minor depression: a population-based study. *Acta Psychiatr Scand* 2007 Mar;115(3):229-236. [doi: [10.1111/j.1600-0447.2006.00851.x](https://doi.org/10.1111/j.1600-0447.2006.00851.x)] [Medline: [17302623](#)]
17. 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](#)]
18. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* 2007 Sep 8;370(9590):851-858. [doi: [10.1016/S0140-6736\(07\)61415-9](https://doi.org/10.1016/S0140-6736(07)61415-9)] [Medline: [17826170](#)]
19. Saxena S, Jané-Llopis E, Hosman C. Prevention of mental and behavioural disorders: implications for policy and practice. *World Psychiatry* 2006 Feb;5(1):5-14. [Medline: [16757984](#)]
20. Birmaher B, Ryan ND, Williamson DE, Brent DA, Kaufman J. Childhood and adolescent depression: a review of the past 10 years. Part II. *J Am Acad Child Adolesc Psychiatry* 1996 Dec;35(12):1575-1583. [doi: [10.1097/00004583-199612000-00008](https://doi.org/10.1097/00004583-199612000-00008)] [Medline: [8973063](#)]
21. Vanheusden K, Mulder CL, van der Ende J, van Lenthe FJ, Mackenbach JP, Verhulst FC. Young adults face major barriers to seeking help from mental health services. *Patient Educ Couns* 2008 Oct;73(1):97-104. [doi: [10.1016/j.pec.2008.05.006](https://doi.org/10.1016/j.pec.2008.05.006)] [Medline: [18584997](#)]
22. Lépine JP, Gastpar M, Mendlewicz J, Tylee A. Depression in the community: the first pan-European study DEPRES (Depression Research in European Society). *Int Clin Psychopharmacol* 1997 Jan;12(1):19-29. [Medline: [9179630](#)]
23. Nuijen J, Achterberg W, van Asch I, van Dijk C, van Dijk S, Frijters D, et al. Trendrapportage GGZ 2010: toegang en zorggebruik [Trend Report on Mental Health Care 2010: Access and Usage of Mental Health Care. Commissioned by the Dutch Ministry of Health, Welfare and Sport]. Utrecht, Netherlands: Trimbos Institute; 2010.
24. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009 Dec;38(4):196-205. [doi: [10.1080/16506070903318960](https://doi.org/10.1080/16506070903318960)] [Medline: [20183695](#)]
25. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](#)]
26. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychol Med* 2010 Dec;40(12):1943-1957. [doi: [10.1017/S0033291710000772](https://doi.org/10.1017/S0033291710000772)] [Medline: [20406528](#)]
27. van der Zanden AP, Conijn BP, Visscher RFM, Gerrits RS. Gripopjedip online: een preventieve groepsinterventie voor jongeren via Internet [Master your mood online: a preventive web-based group course for youth]. *Maandbl Geestelijke Volksgezond* 2005;60(5):515-526.
28. Barak A, Wander-Schwartz M. *Journal of Virtual Environments*. 2000. Empirical Evaluation of Brief Group Therapy Conducted in an Internet Chat Room URL: <http://construct.haifa.ac.il/~azy/cherapy.htm> [accessed 2010-08-24] [WebCite Cache ID 5sDWYfwuf]
29. Gerrits RS, van der Zanden AP, Visscher RFM, Conijn BP. Master your mood online: a preventive chat group intervention for adolescents. *Aust E J Adv Ment Health* 2007;6(3):1-11.
30. van der Zanden RA, Speetjens PA, Arntz KS, Onrust SA. Online group course for parents with mental illness: development and pilot study. *J Med Internet Res* 2010;12(5):e50 [FREE Full text] [doi: [10.2196/jmir.1394](https://doi.org/10.2196/jmir.1394)] [Medline: [21169178](#)]
31. Caele AL, Christensen H. Review of internet-based prevention and treatment programs for anxiety and depression in children and adolescents. *Med J Aust* 2010 Jun 7;192(11 Suppl):S12-S14. [Medline: [20528700](#)]
32. Riper H, Smit F, van der Zanden AP, Conijn B, Kramer J. High tech, high touch, high trust. Programmeringsstudie E-Mental Health in opdracht van het Ministerie van Volksgezondheid, Welzijn en Sport [High Tech, High Touch, High Trust. E-mental Health Programming Study by Order of the Ministry of Health, Welfare and Sport]. Utrecht, Netherlands: Trimbos Institute; 2007.
33. Caele AL, Christensen H, Mackinnon A, Griffiths KM, O'Kearney R. The YouthMood Project: a cluster randomized controlled trial of an online cognitive behavioral program with adolescents. *J Consult Clin Psychol* 2009 Dec;77(6):1021-1032. [doi: [10.1037/a0017391](https://doi.org/10.1037/a0017391)] [Medline: [19968379](#)]

34. Van Voorhees BW, Fogel J, Reinecke MA, Gladstone T, Stuart S, Gollan J, et al. Randomized clinical trial of an Internet-based depression prevention program for adolescents (Project CATCH-IT) in primary care: 12-week outcomes. *J Dev Behav Pediatr* 2009 Feb;30(1):23-37 [FREE Full text] [doi: [10.1097/DBP.0b013e3181966c2a](https://doi.org/10.1097/DBP.0b013e3181966c2a)] [Medline: [19194326](https://pubmed.ncbi.nlm.nih.gov/19194326/)]
35. van der Zanden RA, Kramer JJ, Cuijpers P. Effectiveness of an online group course for adolescents and young adults with depressive symptoms: study protocol for a randomized controlled trial. *Trials* 2011;12:196 [FREE Full text] [doi: [10.1186/1745-6215-12-196](https://doi.org/10.1186/1745-6215-12-196)] [Medline: [21854617](https://pubmed.ncbi.nlm.nih.gov/21854617/)]
36. 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](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
37. van Vliet IM, Leroy H, van Megen H. M.I.N.I. Plus. International Neuropsychological Interview. Nederlandse Versie 5.0.0. [Dutch Version 5.0.0]. Utrecht, Netherlands: UMC; 2000.
38. De Jutters, Indigo, Dimence, and Trimbos Instituut. 2010. Grip op je Dip URL: <http://www.gripopjedip.nl/nl/Home/> [accessed 2012-04-10] [WebCite Cache ID 66oqkHt9j]
39. Cuijpers P, Muñoz RF, Clarke GN, Lewinsohn PM. Psychoeducational treatment and prevention of depression: the "Coping with Depression" course thirty years later. *Clin Psychol Rev* 2009 Jul;29(5):449-458. [doi: [10.1016/j.cpr.2009.04.005](https://doi.org/10.1016/j.cpr.2009.04.005)] [Medline: [19450912](https://pubmed.ncbi.nlm.nih.gov/19450912/)]
40. Lewinsohn PM, Antonucci DO, Breckenridge JS, Teri L. The Coping With Depression Course. Eugene, OR: Castalia Publishing Company; 1984.
41. Gerrits RS, van der Zanden RAP. Gripopjedip online voor jongeren van 16-25 jaar met depressieklachten. Draaiboek voor cursusbegleiders [Master Your Mood Online for Youth Aged 16-25 Having Depressive Symptoms. Course Manual]. Utrecht, Netherlands: Trimbos Institute; 2008.
42. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1(3):385-401.
43. Bouma J, Ranchor AV, Sanderma R, van Sonderen E. Het meten van symptomen van depressie met de CES-D. Een handleiding [Measurement of Symptoms of Depression With the CES-D: A Manual]. Groningen, Netherlands: Rijksuniversiteit Groningen; 1995.
44. Ogles BM, France CR, Lunnen KM, Bell MT, Goldfarb M. Computerized depression screening and awareness. *Community Ment Health J* 1998 Feb;34(1):27-38. [Medline: [9559238](https://pubmed.ncbi.nlm.nih.gov/9559238/)]
45. Cuijpers P, Boluijt P, van Straten A. Screening of depression in adolescents through the Internet : sensitivity and specificity of two screening questionnaires. *Eur Child Adolesc Psychiatry* 2008 Feb;17(1):32-38. [doi: [10.1007/s00787-007-0631-2](https://doi.org/10.1007/s00787-007-0631-2)] [Medline: [17876508](https://pubmed.ncbi.nlm.nih.gov/17876508/)]
46. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [Medline: [6880820](https://pubmed.ncbi.nlm.nih.gov/6880820/)]
47. 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](https://pubmed.ncbi.nlm.nih.gov/9089829/)]
48. Pearlin LI, Schooler C. The structure of coping. *J Health Soc Behav* 1978 Mar;19(1):2-21. [Medline: [649936](https://pubmed.ncbi.nlm.nih.gov/649936/)]
49. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B Stat Methodol* 1977;39(1):1-18.
50. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Erlbaum; 1998.
51. Jacobson NS, Truax P. Clinical significance : a statistical approach to defining meaningful change in psychotherapy research. In: Kazdin AE, editor. *Methodological Issues & Strategies in Clinical Research*. Washington, DC: American Psychological Association; 1992:631-648.
52. Clarke GN, Rohde P, Lewinsohn PM, Hops H, Seeley JR. Cognitive-behavioral treatment of adolescent depression: efficacy of acute group treatment and booster sessions. *J Am Acad Child Adolesc Psychiatry* 1999 Mar;38(3):272-279. [doi: [10.1097/00004583-199903000-00014](https://doi.org/10.1097/00004583-199903000-00014)] [Medline: [10087688](https://pubmed.ncbi.nlm.nih.gov/10087688/)]
53. Michael KD, Crowley SL. How effective are treatments for child and adolescent depression? A meta-analytic review. *Clin Psychol Rev* 2002 Mar;22(2):247-269. [Medline: [11806021](https://pubmed.ncbi.nlm.nih.gov/11806021/)]
54. Reinecke MA, Ryan NE, DuBois DL. Cognitive-behavioral therapy of depression and depressive symptoms during adolescence: a review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 1998 Jan;37(1):26-34. [doi: [10.1097/00004583-199801000-00013](https://doi.org/10.1097/00004583-199801000-00013)] [Medline: [9444896](https://pubmed.ncbi.nlm.nih.gov/9444896/)]
55. Spek V, Cuijpers P, Nyklíček 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](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]
56. Titov N, Dear BF, Schwencke G, Andrews G, Johnston L, Craske MG, et al. Transdiagnostic internet treatment for anxiety and depression: a randomised controlled trial. *Behav Res Ther* 2011 Aug;49(8):441-452. [doi: [10.1016/j.brat.2011.03.007](https://doi.org/10.1016/j.brat.2011.03.007)] [Medline: [21679925](https://pubmed.ncbi.nlm.nih.gov/21679925/)]

57. 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](https://doi.org/10.2196/jmir.954)] [Medline: [18364344](#)]
58. 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](https://doi.org/10.2196/jmir.1094)] [Medline: [19033149](#)]
59. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. *J Med Internet Res* 2005;7(2):e16 [FREE Full text] [doi: [10.2196/jmir.7.2.e16](https://doi.org/10.2196/jmir.7.2.e16)] [Medline: [15998607](#)]
60. Clarke GN, Hornbrook M, Lynch F, Polen M, Gale J, Beardslee W, et al. A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Arch Gen Psychiatry* 2001 Dec;58(12):1127-1134. [Medline: [11735841](#)]
61. Spek V, Nyklíček I, Smits N, Cuijpers P, Riper H, Keyzer J, et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. *Psychol Med* 2007 Dec;37(12):1797-1806. [doi: [10.1017/S0033291707000542](https://doi.org/10.1017/S0033291707000542)] [Medline: [17466110](#)]
62. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychol Bull* 2006 Jan;132(1):132-149. [doi: [10.1037/0033-2909.132.1.132](https://doi.org/10.1037/0033-2909.132.1.132)] [Medline: [16435960](#)]
63. Jensen PS. What is the evidence for evidence-based treatments? *Emot Behav Disord Youth* 2003;3:37-41.
64. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](#)]
65. 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](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](#)]

Abbreviations

CBT: cognitive behavioral therapy
CES-D: Center for Epidemiologic Studies Depression Scale
CI: confidence interval
HADS: Hospital Anxiety and Depression Scale
MINI-Plus: Mini-International Neuropsychiatric Interview
MYM: Master Your Mood

Edited by G Eysenbach; submitted 21.12.11; peer-reviewed by A Calcar; comments to author 28.03.12; revised version received 11.04.12; accepted 23.04.12; published 07.06.12.

Please cite as:

van der Zanden R, Kramer J, Gerrits R, Cuijpers P
Effectiveness of an Online Group Course for Depression in Adolescents and Young Adults: A Randomized Trial
J Med Internet Res 2012;14(3):e86
URL: <http://www.jmir.org/2012/3/e86/>
doi: [10.2196/jmir.2033](https://doi.org/10.2196/jmir.2033)
PMID: [22677437](https://pubmed.ncbi.nlm.nih.gov/22677437/)

©Rianne van der Zanden, Jeannet Kramer, Rob Gerrits, Pim Cuijpers. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 07.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Novel Evaluation of World No Tobacco Day in Latin America

John W Ayers^{1,2,3}, MA, PhD; Benjamin M Althouse⁴, ScM; Jon-Patrick Allem⁵, MA; Daniel E Ford^{4,6}, MPH, MD; Kurt M Ribisl^{7,8}, PhD; Joanna E Cohen⁴, MHSc, PhD

¹Children's Hospital Informatics Program in the Harvard-MIT Division of Health Sciences and Technology, Boston, MA, United States

²Children's Hospital Boston, Boston, MA, United States

³Harvard Medical School, Boston, MA, United States

⁴Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

⁵Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

⁶Johns Hopkins School of Medicine, Baltimore, MD, United States

⁷Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States

⁸Lineberger Comprehensive Cancer Center, Chapel Hill, NC, United States

Corresponding Author:

John W Ayers, MA, PhD

Children's Hospital Informatics Program in the Harvard-MIT Division of Health Sciences and Technology

1 Autumn St

Boston, MA, 02215

United States

Phone: 1 617 355 6000

Fax: 1 617 355 6998

Email: ayers.john.w@gmail.com

Abstract

Background: World No Tobacco Day (WNTD), commemorated annually on May 31, aims to inform the public about tobacco harms. Because tobacco control surveillance is usually annualized, the effectiveness of WNTD remains unexplored into its 25th year.

Objective: To explore the potential of digital surveillance (infoveillance) to evaluate the impacts of WNTD on population awareness of and interest in cessation.

Methods: Health-related news stories and Internet search queries were aggregated to form a continuous and real-time data stream. We monitored daily news coverage of and Internet search queries for cessation in seven Latin American nations from 2006 to 2011.

Results: Cessation news coverage peaked around WNTD, typically increasing 71% (95% confidence interval [CI] 61–81), ranging from 61% in Mexico to 83% in Venezuela. Queries indicative of cessation interest peaked on WNTD, increasing 40% (95% CI 32–48), ranging from 24% in Colombia to 84% in Venezuela. A doubling in cessation news coverage was associated with approximately a 50% increase in cessation queries. To gain a practical perspective, we compared WNTD-related activity with New Year's Day and several cigarette excise tax increases in Mexico. Cessation queries around WNTD were typically greater than New Year's Day and approximated a 2.8% (95% CI –0.8 to 6.3) increase in cigarette excise taxes.

Conclusions: This novel evaluation suggests WNTD had a significant impact on popular awareness (media trends) and individual interest (query trends) in smoking cessation. Because WNTD is constantly evolving, our work is also a model for real-time surveillance and potential improvement in WNTD and similar initiatives.

(*J Med Internet Res* 2012;14(3):e77) doi:[10.2196/jmir.2148](https://doi.org/10.2196/jmir.2148)

KEYWORDS

Tobacco smoking; evaluation research; health communication; informatics; infoveillance; infodemiology

Introduction

Tobacco is responsible for about 4 million premature deaths each year, mostly in developing countries [1] where smoking continues to find safe harbor [2]. Latin America has the seventh and third highest smoking-related mortality among men and women, respectively, in the 14 epidemiologic global regions [1]. Latin America's share of smoking-attributable mortality will likely increase given the high prevalence of smoking: 43% of men and 23% of women smoke in Peru, 41% and 31% in Chile, 37% and 12% in Mexico, and 34% and 24% in Argentina, compared with 23% and 18% in the United States [3,4].

Mandated by the World Health Assembly in 1987 and commemorated annually on May 31, World No Tobacco Day (WNTD) aims to inform the public about the global tobacco epidemic using a combination of mass media approaches, including television, radio, and the Internet [5]. WNTD is believed to be especially important in regions with limited resources for cessation campaigns [6,7] and limited (albeit rapidly changing) tobacco control policy provisions, like in Latin America [8-11]. However, nearing the 25th anniversary of WNTD, its effectiveness remains unexplored.

Because tobacco surveillance is typically annualized, it is exceedingly difficult to estimate the impact of a single awareness day [12]. In nearly all studies [13], investigators make comparisons using annual data from cross-sectional or cohort surveys before and after the intervention [14]; however, differences across the year may not be attributable to the intervention effect [15]. Sometimes investigators ask respondents in cross-sectional surveys to describe their exposure to an intervention, comparing those reporting exposure with those not reporting exposure [16]. Rarely, investigators use cross-sectional monthly survey responses to yield estimates on a finer temporal resolution [17], but these trends are costly to obtain and are still not useful for a single awareness day, given the primary effects are expected on a finer temporal resolution than months. Pioneering work in infodemiology and infoveillance provided a framework for how to assess WNTD's potential effectiveness [18-23]. For instance, Twitter feeds presage cholera epidemics [24]; Internet search queries forecast dengue incidence [25]; and unstructured media reports provide the earliest alerts of disease outbreaks [26]. Building on this work, we employed a novel daily digital surveillance approach using freely available and public archives to estimate the effectiveness of WNTD in seven Latin American nations.

We monitored both intermediate and primary outcomes. First, WNTD may promote cessation awareness through the promotion of smoking cessation in news media [13]. By monitoring online news archives for articles promoting cessation, we captured cessation awareness trends that may have been motivated by WNTD. Second, WNTD may encourage smokers to search the Web for cessation resources [27]. By monitoring aggregate Internet search queries indicative of cessation, we captured interest in cessation that may have been motivated by WNTD. WNTD increases in cessation awareness may trigger increases in cessation interest; therefore, we examined the relationship

between cessation news coverage and Internet search queries for cessation.

Methods

To address the study objectives, we analyzed online news media promoting cessation and online search queries for cessation from 2006 to 2011 in Mexico, Colombia, Argentina, Peru, Venezuela, Chile, and Ecuador using a quasi-experimental design [15].

We monitored weekly news stories archived on Google News that mentioned quitting smoking (Google Inc, Mountain View, CA, USA; news.google.com). Google News captures a broad spectrum of print, radio, and television media, and therefore is a useful archive to monitor cessation media stories. To identify the country of origin for each article, we downloaded data from nation-specific Google News domains (eg, news.google.com.mx for Mexico). All stories that mentioned smoking cessation were monitored relative to all stories in each nation. The numerator was the number of articles containing the phrase “dejar de fumar” (literally “give up smoking”), chosen based on discussion with experts and native Spanish speakers familiar with linguistic nuances across the nations monitored. The denominator was the number of stories that contained any universal conjunctive phrases, such as those containing “de,” thereby indicating a count of *all* stories. We report the resulting quotient per 100,000 stories.

Aggregate query trends were downloaded from Google Insights for Search (www.google.com/insights/). We monitored all queries that included “dejar de fumar” by nation. Trends were analyzed on a relative search volume (RSV) scale, with queries normalized to the period with the highest search proportion—for example, $RSV = 100$ is the highest search proportion period (cessation queries versus all search queries), and $RSV = 50$ is 50% of the highest search proportion. This approach corrects for trending in absolute search volume that is usually increasing for all common queries [28].

We analyzed data on weekly and daily time trends. Visual inspection of the data suggested the likely impact of WNTD was a pulse effect, an immediate change in mean news or query trends, rather than a durable shift. Because WNTD is a global program, there are no nations that may serve as untreated controls. As a result, we specified interrupted time series [29] within each nation to estimate the mean RSV around and including WNTD (3 weeks, including the week of, the week before, and the week after WNTD) to capture the rise and decline in WNTD spikes in reference to a within-nation RSV sampling from the 12 weeks before and after the WNTD period. The weeks before and after WNTD were included to capture presaging and lingering increases in news and search that was visually evident in the data inspection. We estimated the effect size as a ratio between WNTD and reference periods to make estimates comparable across nations ($Ho: (RSV_{WNTD} - RSV_{reference\ weeks}) / RSV_{reference\ weeks} = 0$) [30]. To find the amount of variance in search queries explained by media trends, we regressed news trends lagged by 0, 1, and 2 weeks against search

terms and reported r^2 for all periods and after selecting only periods around WNTD [31].

Results

Figure 1 shows example cessation news coverage and cessation query trends in Mexico (patterns were similar in the six other study nations). Both news and query trends spiked in Mexico around WNTD. Averaging across years, cessation news trends typically peaked at 30 per 100,000 news stories the week of WNTD with cessation query trends peaking at 95% (RSV) of the highest search proportion week (with the highest RSV, 100, being on New Year's Day 2008). Average trends across years (Figure 1B) suggested increases around WNTD were greater than the usual New Year's spikes in Mexico. News and query peaks varied in their timing where there were strong buildups in news coverage beginning almost a month before WNTD, terminating 2 or 3 days after WNTD (Figure 1C), and queries had a pronounced peak around WNTD with queries remaining higher for a week after WNTD. Cessation news trends averaged 147 per 100,000 stories on WNTD, and queries had the highest search proportion day (RSV = 100) on WNTD.

Statistical analysis suggested WNTD was associated with increases in cessation news coverage and cessation queries in all nations (Figure 2). Averaging across all nations and years suggested cessation news trends were 71% (95% confidence interval [CI] 61–81) higher and cessation query trends were 40% (95% CI 32–48) higher around WNTD than expected, as compared with trends 12 weeks before and after the WNTD period. Cessation news increases varied year to year, ranging from 24% (95% CI 18–30) to 116% (95% CI 103–128) for 2006 and 2009, respectively. Cessation query increases ranged from 28% (95% CI 21–36) to 50% (95% CI 43–57) for 2009 and 2008, respectively. Analyses by nation suggested news mentions of cessation ranged from 61% (95% CI 31–91) to 83% (95% CI 53–113) higher around WNTD in Mexico and Venezuela, respectively (Figure 3). Cessation queries ranged from 24% (95% CI 15–32) to 84% (95% CI 68–100) higher around WNTD in Argentina and Venezuela, respectively.

Cessation news coverage explained limited amounts of variance in cessation queries. Multiple r^2 ranged from .024 in Argentina to .516 in Mexico. Generally, cessation news coverage and cessation queries were not strongly associated across all time periods. However, the correlation between mean trends 3 months before to 3 months after WNTD was substantially higher (eg, $r = .5$ in Mexico). Figure 4 compares the magnitude of the effect sizes for search and news trend increases around WNTD by nation and year. Analyses showed increases in news trends (the week of, and before and after WNTD) were larger than increases in search query trends for nearly all the years, suggesting increases in cessation news coverage had diminishing returns on cessation queries during this narrow period. As such, a doubling in cessation news coverage corresponded with about a 50% increase in cessation queries.

Increases in cessation news coverage and cessation queries around WNTD need to be placed in practical perspective. Cigarette excise tax increases are the most effective tobacco control policy [32]. Around New Year's Day in 2007, 2008, 2009, and 2011, Mexico increased its cigarette excise tax 4.7%, 1.3%, 1.2%, and 7.8%, respectively. We compared increases in cessation news coverage and cessation queries around WNTD with those from these tax years. In Mexico, increases in cessation queries around New Year's Day (compared with the WNTD baseline) were as little as 10% (95% CI 2–23). WNTD-related increases were typically much larger than New Year's-related increases in absence of a tax increase and were sometimes greater than increases when a new tax increase was also implemented. Regressing tax increases against WNTD increases across years and predicting the result for the overall mean suggested that WNTD increases in cessation news coverage and query trends approximated those for a 2.8% (95% CI –0.8 to 6.3) cigarette excise tax increase. Additionally, the estimate is based on a small sample (6 years), so while it is qualitatively accurate, the effect of WNTD may be equivalent to a somewhat lower or higher tax, especially since all tax increases occurred in addition to any New Year's effects.

Figure 1. Example World No Tobacco Day (WNTD) time trends in Mexico, 2006-2011. (A) Entire series by week, (B) entire weekly series with mean annualized estimate, and (C) daily series with mean estimates. Daily query estimates were restricted to 2010 and 2011 due to data unavailability.

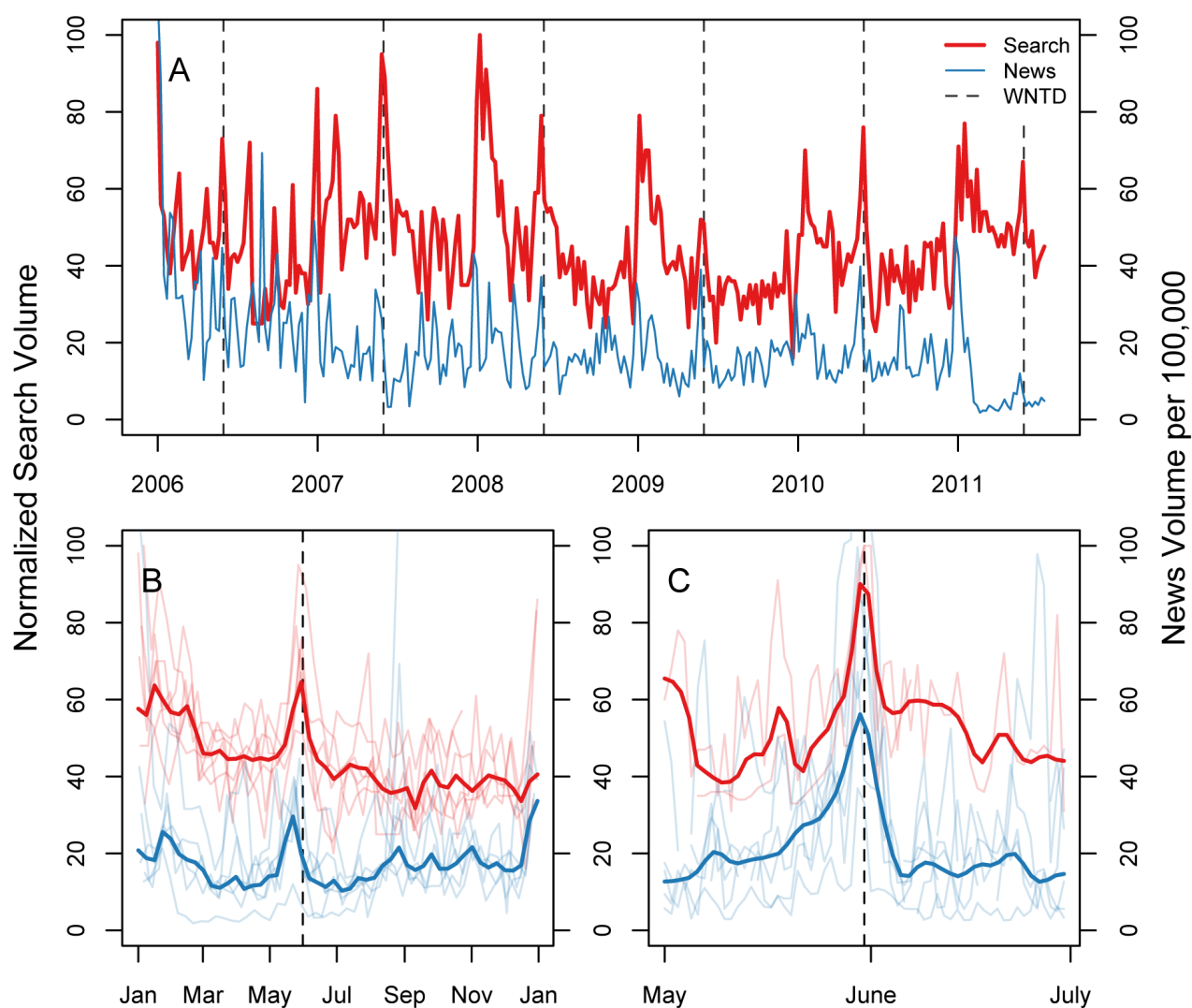


Figure 2. Pooled estimates of World No Tobacco Day effectiveness in Latin America, 2006-2011. Estimates are made comparing the week of (and before and after) World No Tobacco Day with 12 weeks before and after that period.

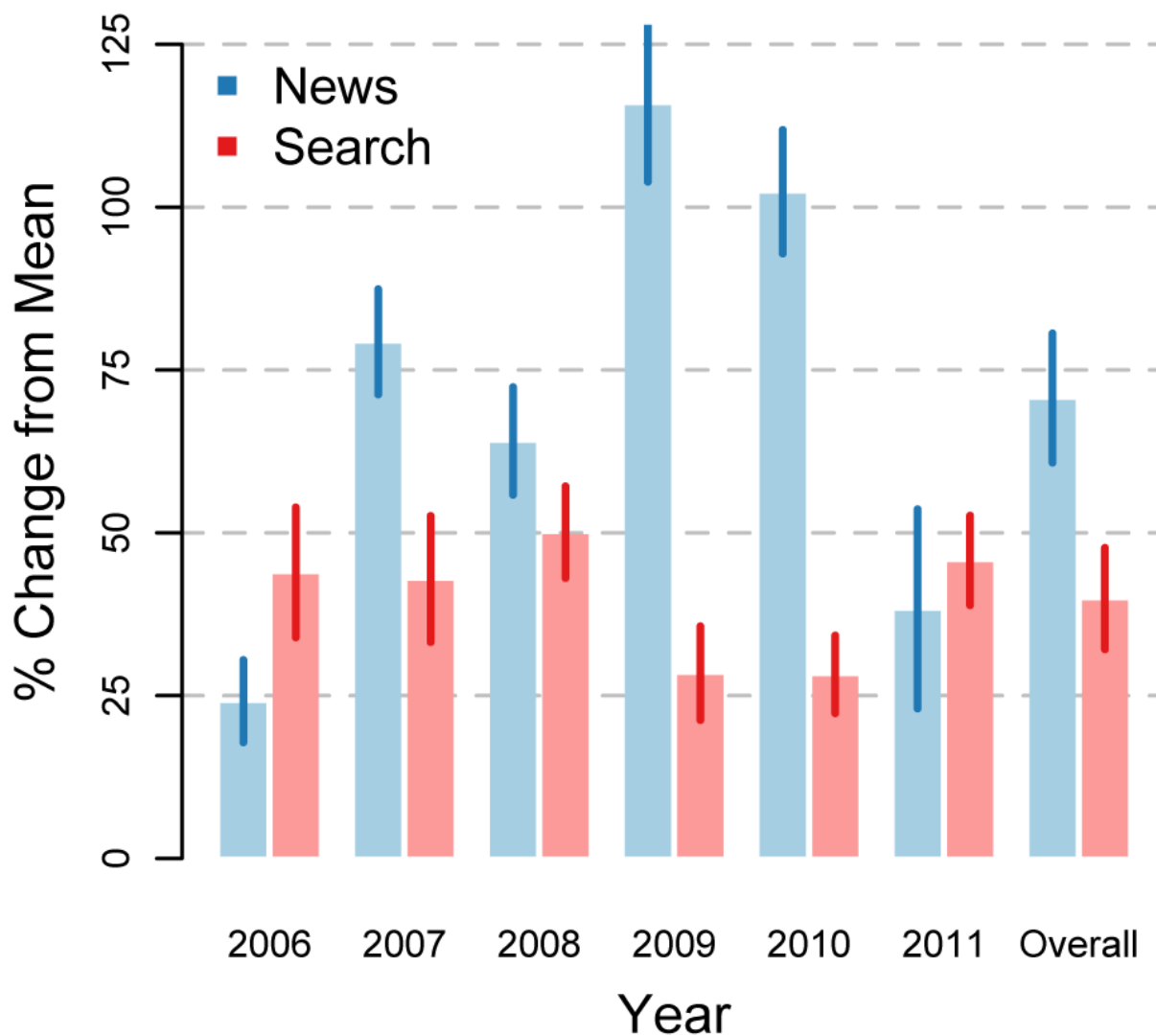


Figure 3. Nation-specific estimates of World No Tobacco Day effectiveness in Latin America, 2006-2011. Dotted line indicates average effect for all nations and years as a reference point. Data were unavailable in Ecuador, Chile, and Colombia for 2006 and 2007.

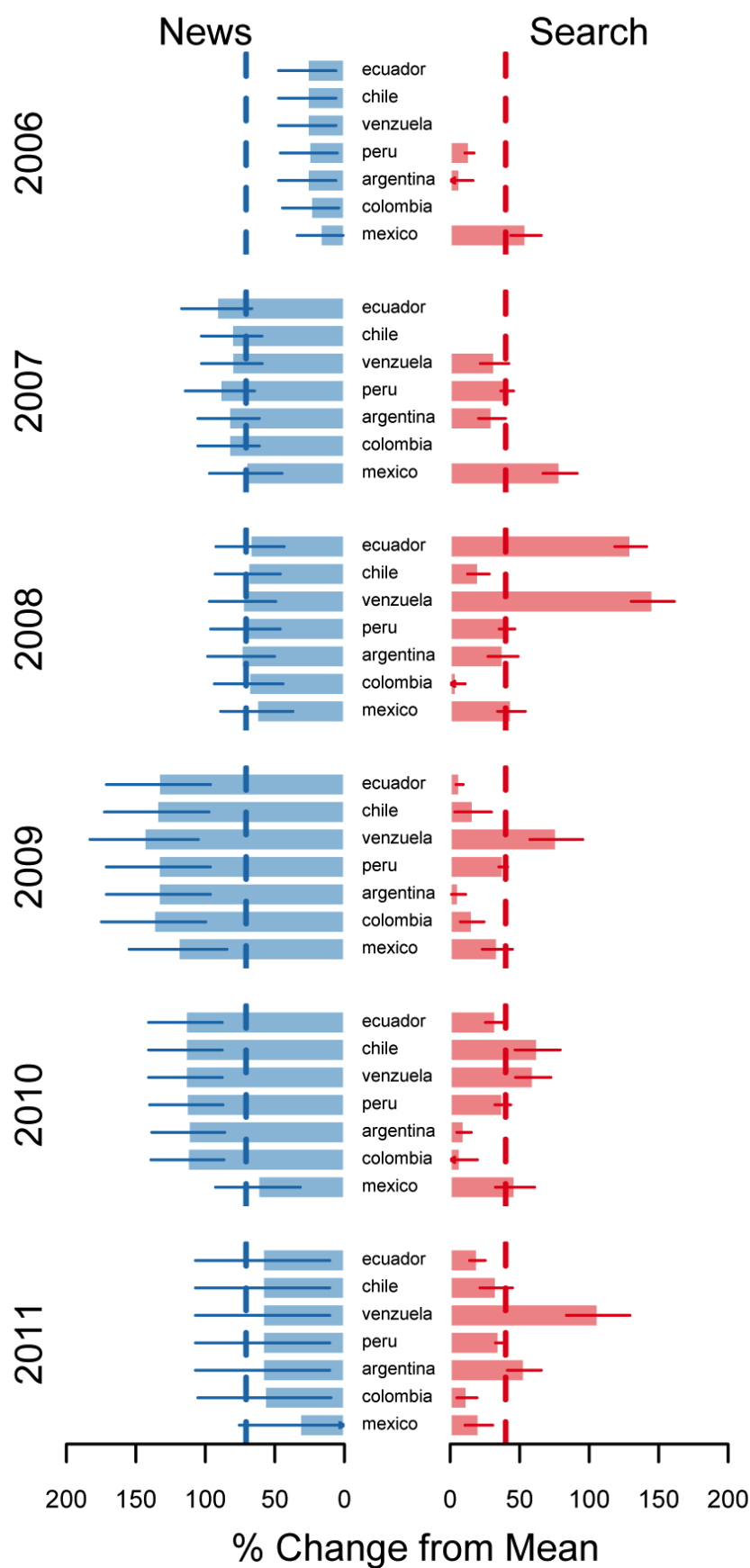
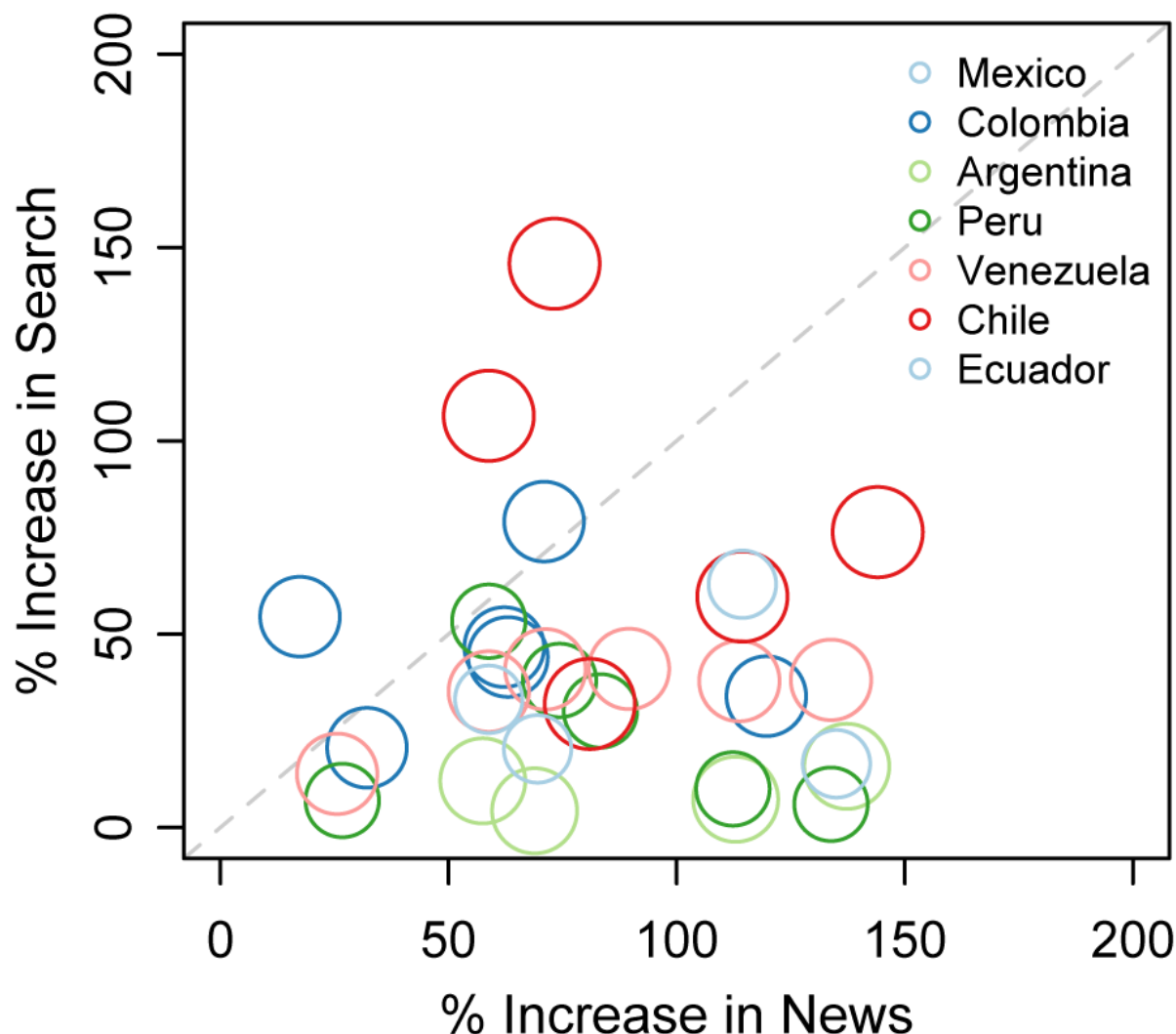


Figure 4. Cessation news coverage is positively associated with cessation Internet search queries. Nodes are sized according to mean annual search query volume, with nodes by nation and year.



Discussion

We demonstrated the feasibility of novel digital surveillance for tobacco control, yielding the first population estimates of WNTD's effectiveness in the first program evaluation to use digital surveillance. WNTD coincided with large increases in cessation news coverage and cessation Internet search queries in seven Latin American nations. WNTD increases eclipsed New Year's Day increases and approximated the effectiveness of the combination of New Year's and a 2.8% cigarette tax hike. There was little correlation between news trends and query trends when comparing across the entire time series, but higher correlation when comparing trends around WNTD. This may indicate that news plays a minor role in increasing cessation queries generally but is magnified when the population is primed by WNTD [33].

This report builds on a programmatic agenda complementing tobacco control surveillance with digital data streams [34,35]. Epidemiologists have demonstrated the potential of digital surveillance for monitoring chickenpox [36,37], dengue [25,38],

gastritis [36], influenza [18,39] including H1N1 [20], kidney stones [40,41], listeriosis [23], Lyme disease [42], methicillin-resistant *Staphylococcus aureus* [43], and salmonella outbreaks [22]. However, applications to health behaviors are very rare [20,27,34,35,44]. With smoking accounting for almost 6 million deaths each year [45], the application of digital surveillance to tobacco control may be of great significance. Herein we present our approach as a viable method to evaluate WNTD where traditional methods were ill equipped to do so.

While the infoveillance approach has many strengths, it is not without limitations. We only considered increase in news coverage around WNTD and did not consider the placement of news coverage (eg, lead versus minor story), but this may be captured by news volume, where increased volume is associated with premium placement. Cessation interest was only assessed for smokers with access to the Internet, but we assume that changes in cessation interest among those with access mirrors those without access [46]. Recent studies encourage this assumption, suggesting queries are valid proxies for a host of health outcomes, even in disparate regions [25,39]. For example,

depression queries correlate with suicide rates [47] and cancer queries correlate with cancer-specific incidence rates [48]. We compared WNTD with cigarette excise taxes in Mexico (the only nation with several cigarette excise tax increases), and all tax increases occurred on New Year's Day; hence, WNTD's approximation to a 2.8% tax increase might be larger in other nations and when taxes do not increase during New Year's.

Global health awareness campaigns are difficult to evaluate given sparse transnational data streams, and contemporary surveillance does not collect data on the same time dimensions as the campaigns that will likely effect change (eg, annual surveillance versus daily impacts) [12]. Tobacco control can learn much from adapting novel techniques that are gaining traction in biosurveillance [38]. Here we have built on validated approaches used in disease surveillance [25,38] and highlighted the potential advantages of these techniques over other tobacco control surveillance methods. Survey-based tobacco control surveillance has high financial costs, lacks timeliness, and may not reach some populations (eg, cell phone-only households) [49]. Also, responses gathered may be biased by strong social desirability [50]. Alternatives using passive data generation are only reliable in certain geographic areas and still require costly data processing (eg, nicotine replacement therapy sales) [51]. Digital surveillance, as such, is a critical forefront for tobacco control with real-time, public, and low-cost data streams useful for program and policy evaluations. Analyses of media and search query trends also afford greater transparency, as scientists may quickly replicate each other's work, downloading data from regularly updated online archives. Extending our approach can serve as a continual evaluation system for global WNTD effects.

In part, these results may outline how well WNTD's yearly themes resonate with Latin American populations. In 2009 and 2010 when WNTD focused on tobacco health warnings and smoking among women, respectively, cessation news coverage was the highest. News coverage increases were smallest when WNTD focused on general campaigns, such as promoting the World Health Organization Framework Convention on Tobacco Control (2011) and promoting awareness of the health effects of nonsmoking-related tobacco harms (2006). The variability

in cessation search queries was much smaller year to year, but increases were generally larger during WNTD's clean indoor air (2007) and restricting youth access to tobacco (2008) campaigns.

WNTD planners, and those responsible for tobacco cessation initiatives more broadly, may also use our results and techniques to scale up and refine the online components of their campaigns [52]. Since many search engines rank links according to the number of times searchers visit the page, or how many other pages link to a page, it is unrealistic to assume that creating cessation webpages, blogs, or YouTube videos will necessarily reach large numbers of tobacco users, even if they are seeking these resources. Search engine advertisements triggered by user-specific queries may be purchased for as little as US \$0.01 per click. Advertised links appear on the first page of search results, ensuring that cessation queries meet webpages with current, health-conscious, and objective discussion of cessation when program planners advertise online. Search queries are a rare opportunity to reach smokers when they are thinking about cessation, while at the same time taking into account their stage in the cessation process with links individually tailored to optimize effectiveness [53,54]. By appropriately purchasing ad links, planners can ensure that evidence-based approaches maintain a strong first-page presence in cessation search results, especially around WNTD when many smokers appear interested in cessation. This is especially crucial given that most smokers attempt to quit without professional counseling or pharmacologic therapy [55].

The effectiveness of WNTD has been widely speculated [8], but to our knowledge these are the first empirical estimates of its impact. In evaluating this intervention, we highlight the importance of digital surveillance for performing difficult evaluations using real-time, low-cost, and transparent methods. Herein we find that WNTD is producing large increases in cessation news coverage and cessation Internet search queries, rivaling the effects of Mexico's recent tobacco tax increases and New Year's resolutions. Because WNTD is constantly evolving, our work is also a framework for real-time surveillance and potential improvement in WNTD.

Acknowledgments

This research was supported by a grant to JWA from the Institute for Global Tobacco Control with funding from the Bloomberg Initiative to Reduce Tobacco Use. JWA was also supported by a grant from Google.org and a National Institute of Child Health & Human Development (NICHD) Training Grant in Pediatric Emergency Medicine (5T32HD040128). A National Science Foundation (NSF) Graduate Research Fellowship also supported BMA. The Bloomberg Initiative to Reduce Tobacco Use, Google.org, NICHD, and NSF had no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript. This report was improved by comments from a colloquium at Children's Hospital Boston, University of North Carolina, and University of Southern California Keck School of Medicine. We thank John Brownstein, Elizabeth Kennedy, Ken Mandl, Seth Noar, Steve Sussman, and Jennifer Unger for their helpful criticisms.

Authors' Contributions

Study concept and design: JWA, JEC. Acquisition of data: JWA, BMA. Analysis and interpretation of data: JWA, BMA, JEC. Drafting of the manuscript: JWA, BMA, JPA. Critical revision of the manuscript for important intellectual content: JWA, BMA, JPA, DEF, KMR, JEC.

Conflicts of Interest

JWA and BMA share an equity stake in a consulting group, Directing Medicine, that helps others implement some of the ideas embodied in this work. The data generation procedures, however, are not proprietary and rely on public archives. There are no other conflicts of interest relevant to this study.

Editorial note: This paper was originally peer-reviewed by another journal. JMIR reviewed the paper internally and accepted the peer-review reports and the authors' responses as forwarded by the author. Thus, the identity of the reviewers is not known to the editor.

References

1. Ezzati M, Lopez AD. Regional, disease specific patterns of smoking-attributable mortality in 2000. *Tob Control* 2004 Dec;13(4):388-395 [FREE Full text] [doi: [10.1136/tc.2003.005215](https://doi.org/10.1136/tc.2003.005215)] [Medline: [15564623](https://pubmed.ncbi.nlm.nih.gov/15564623/)]
2. Pampel F. National Income, Inequality and Global Patterns of Cigarette Use. *Soc Forces* 2007 Dec 1;86(2):445-466. [Medline: [21874072](https://pubmed.ncbi.nlm.nih.gov/21874072/)]
3. World Lung Foundation, American Cancer Society. The Tobacco Atlas URL: <http://www.tobaccoatlas.org/> [accessed 2012-05-17] [WebCite Cache ID 67k0IySdb]
4. Centers for Disease Control and Prevention (CDC). Cigarette smoking among adults and trends in smoking cessation - United States, 2008. *MMWR Morb Mortal Wkly Rep* 2009 Nov 13;58(44):1227-1232 [FREE Full text] [Medline: [19910909](https://pubmed.ncbi.nlm.nih.gov/19910909/)]
5. World Health Organization. 2012 May 31. World No Tobacco Day URL: <http://www.who.int/mediacentre/events/annual/wntd/en/index.html> [accessed 2012-05-17] [WebCite Cache ID 67k0KX2Px]
6. Müller F, Wehbe L. Smoking and smoking cessation in Latin America: a review of the current situation and available treatments. *Int J Chron Obstruct Pulmon Dis* 2008;3(2):285-293 [FREE Full text] [Medline: [18686737](https://pubmed.ncbi.nlm.nih.gov/18686737/)]
7. Franco-Marina F. Adult smoking trends in Mexico: an analysis of the Mexican National Addiction Surveys. *Salud Publica Mex* 2007;49(suppl 2):s137-s146. [doi: [10.1590/S0036-36342007000800004](https://doi.org/10.1590/S0036-36342007000800004)]
8. Navas-Acien A, Peruga A, Breyse P, Zavaleta A, Blanco-Marquizo A, Pitarque R, et al. Secondhand tobacco smoke in public places in Latin America, 2002-2003. *JAMA* 2004 Jun 9;291(22):2741-2745. [doi: [10.1001/jama.291.22.2741](https://doi.org/10.1001/jama.291.22.2741)] [Medline: [15187056](https://pubmed.ncbi.nlm.nih.gov/15187056/)]
9. Sebríe EM, Blanco A, Glantz SA. Cigarette labeling policies in Latin America and the Caribbean: progress and obstacles. *Salud Publica Mex* 2010;52 Suppl 2:S233-S243 [FREE Full text] [Medline: [21243194](https://pubmed.ncbi.nlm.nih.gov/21243194/)]
10. Thrasher JF, Boado M, Sebríe EM, Bianco E. Smoke-free policies and the social acceptability of smoking in Uruguay and Mexico: findings from the International Tobacco Control Policy Evaluation Project. *Nicotine Tob Res* 2009 Jun;11(6):591-599. [doi: [10.1093/ntr/ntp039](https://doi.org/10.1093/ntr/ntp039)] [Medline: [19380383](https://pubmed.ncbi.nlm.nih.gov/19380383/)]
11. Thrasher JF, Pérez-Hernández R, Swayampakala K, Arillo-Santillán E, Bottai M. Policy support, norms, and secondhand smoke exposure before and after implementation of a comprehensive smoke-free law in Mexico city. *Am J Public Health* 2010 Sep;100(9):1789-1798. [doi: [10.2105/AJPH.2009.180950](https://doi.org/10.2105/AJPH.2009.180950)] [Medline: [20466952](https://pubmed.ncbi.nlm.nih.gov/20466952/)]
12. Castillo-Salgado C. Trends and directions of global public health surveillance. *Epidemiol Rev* 2010 Apr;32(1):93-109. [doi: [10.1093/epirev/mxq008](https://doi.org/10.1093/epirev/mxq008)] [Medline: [20534776](https://pubmed.ncbi.nlm.nih.gov/20534776/)]
13. Bala M, Strzeszynski L, Cahill K. Mass media interventions for smoking cessation in adults. *Cochrane Database Syst Rev* 2008(1):CD004704. [doi: [10.1002/14651858.CD004704.pub2](https://doi.org/10.1002/14651858.CD004704.pub2)] [Medline: [18254058](https://pubmed.ncbi.nlm.nih.gov/18254058/)]
14. Norman GJ, Ribisl KM, Howard-Pitney B, Howard KA, Unger JB. The relationship between home smoking bans and exposure to state tobacco control efforts and smoking behaviors. *Am J Health Promot* 2000;15(2):81-88. [Medline: [11194699](https://pubmed.ncbi.nlm.nih.gov/11194699/)]
15. Shadish WR, Cook TD. The renaissance of field experimentation in evaluating interventions. *Annu Rev Psychol* 2009;60:607-629. [doi: [10.1146/annurev.psych.60.110707.163544](https://doi.org/10.1146/annurev.psych.60.110707.163544)] [Medline: [18616391](https://pubmed.ncbi.nlm.nih.gov/18616391/)]
16. Popham WJ, Potter LD, Bal DG, Johnson MD, Duerr JM, Quinn V. Do anti-smoking media campaigns help smokers quit? *Public Health Rep* 1993;108(4):510-513. [Medline: [8341788](https://pubmed.ncbi.nlm.nih.gov/8341788/)]
17. Hahn EJ, Rayens MK, Butler KM, Zhang M, Durbin E, Steinke D. Smoke-free laws and adult smoking prevalence. *Prev Med* 2008 Aug;47(2):206-209. [doi: [10.1016/j.ypmed.2008.04.009](https://doi.org/10.1016/j.ypmed.2008.04.009)] [Medline: [18519154](https://pubmed.ncbi.nlm.nih.gov/18519154/)]
18. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu Symp Proc* 2006:244-248. [Medline: [17238340](https://pubmed.ncbi.nlm.nih.gov/17238340/)]
19. 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](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
20. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](https://doi.org/10.1016/j.amepre.2011.02.006)] [Medline: [21521589](https://pubmed.ncbi.nlm.nih.gov/21521589/)]
21. Brownstein JS, Freifeld CC, Reis BY, Mandl KD. Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Med* 2008 Jul 8;5(7):e151 [FREE Full text] [doi: [10.1371/journal.pmed.0050151](https://doi.org/10.1371/journal.pmed.0050151)] [Medline: [18613747](https://pubmed.ncbi.nlm.nih.gov/18613747/)]
22. 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](https://doi.org/10.1056/NEJMp0900702)] [Medline: [19423867](https://pubmed.ncbi.nlm.nih.gov/19423867/)]

23. Wilson K, Brownstein JS. Early detection of disease outbreaks using the Internet. *CMAJ* 2009 Apr 14;180(8):829-831. [doi: [10.1503/cmaj.090215](https://doi.org/10.1503/cmaj.090215)] [Medline: [19364791](https://pubmed.ncbi.nlm.nih.gov/19364791/)]
24. Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *Am J Trop Med Hyg* 2012 Jan;86(1):39-45. [doi: [10.4269/ajtmh.2012.11-0597](https://doi.org/10.4269/ajtmh.2012.11-0597)] [Medline: [22232449](https://pubmed.ncbi.nlm.nih.gov/22232449/)]
25. Althouse BM, Ng YY, Cummings DA. Prediction of dengue incidence using search query surveillance. *PLoS Negl Trop Dis* 2011 Aug;5(8):e1258 [FREE Full text] [doi: [10.1371/journal.pntd.0001258](https://doi.org/10.1371/journal.pntd.0001258)] [Medline: [21829744](https://pubmed.ncbi.nlm.nih.gov/21829744/)]
26. Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, Mawudeku A, et al. Use of unstructured event-based reports for global infectious disease surveillance. *Emerg Infect Dis* 2009 May;15(5):689-695 [FREE Full text] [doi: [10.3201/eid1505.081114](https://doi.org/10.3201/eid1505.081114)] [Medline: [19402953](https://pubmed.ncbi.nlm.nih.gov/19402953/)]
27. Cobb NK. Online consumer search strategies for smoking-cessation information. *Am J Prev Med* 2010 Mar;38(3 Suppl):S429-S432. [doi: [10.1016/j.amepre.2009.12.001](https://doi.org/10.1016/j.amepre.2009.12.001)] [Medline: [20176318](https://pubmed.ncbi.nlm.nih.gov/20176318/)]
28. Dutka AF, Hanson HH. *Fundamentals of Data Normalization*. Reading, MA: Addison-Wesley; 1989.
29. Kirchgassner G, Wolters J. *Introduction to Modern Time Series Analysis*. Berlin: Springer; 2007.
30. King G, Tomz M, Wittenberg J. Making the most of statistical analyses: improving interpretation and presentation. *Am J Pol Sci* 2000;44:341-355.
31. Conover WJ. *Practical Nonparametric Statistics*. New York, NY: Wiley; 1999.
32. Jha P, Chaloupka FJ, Corrao M, Jacob B. Reducing the burden of smoking world-wide: effectiveness of interventions and their coverage. *Drug Alcohol Rev* 2006;25:597-609. [Medline: [17132576](https://pubmed.ncbi.nlm.nih.gov/17132576/)]
33. Scheufele DA, Tewksbury D. Framing, agenda setting, and priming: the evolution of three media effects models. *Journal Commun* 2007;57(1):9-20. [doi: [10.1111/j.0021-9916.2007.00326.x](https://doi.org/10.1111/j.0021-9916.2007.00326.x)]
34. Ayers JW, Ribisl K, Brownstein JS. Using search query surveillance to monitor tax avoidance and smoking cessation following the United States' 2009 "SCHIP" cigarette tax increase. *PLoS One* 2011;6(3):e16777 [FREE Full text] [doi: [10.1371/journal.pone.0016777](https://doi.org/10.1371/journal.pone.0016777)] [Medline: [21436883](https://pubmed.ncbi.nlm.nih.gov/21436883/)]
35. Ayers JW, Ribisl KM, Brownstein JS. Tracking the rise in popularity of electronic nicotine delivery systems (electronic cigarettes) using search query surveillance. *Am J Prev Med* 2011 Apr;40(4):448-453. [doi: [10.1016/j.amepre.2010.12.007](https://doi.org/10.1016/j.amepre.2010.12.007)] [Medline: [21406279](https://pubmed.ncbi.nlm.nih.gov/21406279/)]
36. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron A. More diseases tracked by using Google Trends. *Emerg Infect Dis* 2009 Aug;15(8):1327-1328 [FREE Full text] [doi: [10.3201/eid1508.090299](https://doi.org/10.3201/eid1508.090299)] [Medline: [19751610](https://pubmed.ncbi.nlm.nih.gov/19751610/)]
37. Valdivia A, Monge-Corella S. Diseases tracked by using Google trends, Spain. *Emerg Infect Dis* 2010 Jan;16(1):168 [FREE Full text] [doi: [10.3201/eid1601.091308](https://doi.org/10.3201/eid1601.091308)] [Medline: [20031078](https://pubmed.ncbi.nlm.nih.gov/20031078/)]
38. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis* 2011 May;5(5):e1206 [FREE Full text] [doi: [10.1371/journal.pntd.0001206](https://doi.org/10.1371/journal.pntd.0001206)] [Medline: [21647308](https://pubmed.ncbi.nlm.nih.gov/21647308/)]
39. 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](https://doi.org/10.1038/nature07634)] [Medline: [19020500](https://pubmed.ncbi.nlm.nih.gov/19020500/)]
40. Breyer BN, Sen S, Aaronson DS, Stoller ML, Erickson BA, Eisenberg ML. Use of Google Insights for Search to track seasonal and geographic kidney stone incidence in the United States. *Urology* 2011 Aug;78(2):267-271. [doi: [10.1016/j.urology.2011.01.010](https://doi.org/10.1016/j.urology.2011.01.010)] [Medline: [21459414](https://pubmed.ncbi.nlm.nih.gov/21459414/)]
41. Willard SD, Nguyen MM. Internet Search Trends Analysis Tools Can Provide Real-time Data on Kidney Stone Disease in the United States. *Urology* 2011 Jun 13 (forthcoming). [doi: [10.1016/j.urology.2011.04.024](https://doi.org/10.1016/j.urology.2011.04.024)] [Medline: [21676450](https://pubmed.ncbi.nlm.nih.gov/21676450/)]
42. Seifter A, Schwarzwald A, Geis K, Aucott J. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospat Health* 2010 May;4(2):135-137 [FREE Full text] [Medline: [20503183](https://pubmed.ncbi.nlm.nih.gov/20503183/)]
43. Dukic VM, David MZ, Lauderdale DS. Internet queries and methicillin-resistant *Staphylococcus aureus* surveillance. *Emerg Infect Dis* 2011 Jun;17(6):1068-1070 [FREE Full text] [doi: [10.3201/eid1706.101451](https://doi.org/10.3201/eid1706.101451)] [Medline: [21749772](https://pubmed.ncbi.nlm.nih.gov/21749772/)]
44. Reis BY, Brownstein JS. Measuring the impact of health policies using Internet search patterns: the case of abortion. *BMC Public Health* 2010;10:514 [FREE Full text] [doi: [10.1186/1471-2458-10-514](https://doi.org/10.1186/1471-2458-10-514)] [Medline: [20738850](https://pubmed.ncbi.nlm.nih.gov/20738850/)]
45. Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet* 2003 Sep 13;362(9387):847-852. [doi: [10.1016/S0140-6736\(03\)14338-3](https://doi.org/10.1016/S0140-6736(03)14338-3)] [Medline: [13678970](https://pubmed.ncbi.nlm.nih.gov/13678970/)]
46. Stoddard JL, Augustson EM. Smokers who use internet and smokers who don't: data from the Health Information and National Trends Survey (HINTS). *Nicotine Tob Res* 2006 Dec;8 Suppl 1:S77-S85. [Medline: [17491174](https://pubmed.ncbi.nlm.nih.gov/17491174/)]
47. Yang AC, Tsai SJ, Huang NE, Peng CK. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004-2009. *J Affect Disord* 2011 Jul;132(1-2):179-184. [doi: [10.1016/j.jad.2011.01.019](https://doi.org/10.1016/j.jad.2011.01.019)] [Medline: [21371755](https://pubmed.ncbi.nlm.nih.gov/21371755/)]
48. Cooper CP, Mallon KP, Leadbetter S, Pollack LA, Peipins LA. Cancer Internet search activity on a major search engine, United States 2001-2003. *J Med Internet Res* 2005 Jul 1;7(3):e36 [FREE Full text] [doi: [10.2196/jmir.7.3.e36](https://doi.org/10.2196/jmir.7.3.e36)] [Medline: [15998627](https://pubmed.ncbi.nlm.nih.gov/15998627/)]
49. Boland M, Sweeney MR, Scallan E, Harrington M, Staines A. Emerging advantages and drawbacks of telephone surveying in public health research in Ireland and the U.K. *BMC Public Health* 2006;6:208 [FREE Full text] [doi: [10.1186/1471-2458-6-208](https://doi.org/10.1186/1471-2458-6-208)] [Medline: [16911771](https://pubmed.ncbi.nlm.nih.gov/16911771/)]

50. Ayers JW. Measuring English proficiency and language preference: are self-reports valid? *Am J Public Health* 2010 Aug;100(8):1364-6; author reply 1366. [doi: [10.2105/AJPH.2010.194167](https://doi.org/10.2105/AJPH.2010.194167)] [Medline: [20558784](https://pubmed.ncbi.nlm.nih.gov/20558784/)]
51. Metzger KB, Mostashari F, Kerker BD. Use of pharmacy data to evaluate smoking regulations' impact on sales of nicotine replacement therapies in New York City. *Am J Public Health* 2005 Jun;95(6):1050-1055. [doi: [10.2105/AJPH.2004.048025](https://doi.org/10.2105/AJPH.2004.048025)] [Medline: [15914832](https://pubmed.ncbi.nlm.nih.gov/15914832/)]
52. 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](https://doi.org/10.1146/annurev.publhealth.031308.100235)] [Medline: [19296777](https://pubmed.ncbi.nlm.nih.gov/19296777/)]
53. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007 Jul;133(4):673-693. [doi: [10.1037/0033-2909.133.4.673](https://doi.org/10.1037/0033-2909.133.4.673)] [Medline: [17592961](https://pubmed.ncbi.nlm.nih.gov/17592961/)]
54. Herzog TA, Blagg CO. Are most precontemplators contemplating smoking cessation? Assessing the validity of the stages of change. *Health Psychol* 2007 Mar;26(2):222-231. [doi: [10.1037/0278-6133.26.2.222](https://doi.org/10.1037/0278-6133.26.2.222)] [Medline: [17385974](https://pubmed.ncbi.nlm.nih.gov/17385974/)]
55. Chapman S, MacKenzie R. The global research neglect of unassisted smoking cessation: causes and consequences. *PLoS Med* 2010 Feb;7(2):e1000216 [FREE Full text] [doi: [10.1371/journal.pmed.1000216](https://doi.org/10.1371/journal.pmed.1000216)] [Medline: [20161722](https://pubmed.ncbi.nlm.nih.gov/20161722/)]

Abbreviations

CI: confidence interval

RSV: relative search volume

WNTD: World No Tobacco Day

Edited by G Eysenbach; submitted 26.04.12; peer-reviewed by Anonymous, Anonymous, Anonymous; comments to author 14.05.12; revised version received 16.05.12; accepted 16.05.12; published 28.05.12.

Please cite as:

Ayers JW, Althouse BM, Allem JP, Ford DE, Ribisl KM, Cohen JE

A Novel Evaluation of World No Tobacco Day in Latin America

J Med Internet Res 2012;14(3):e77

URL: <http://www.jmir.org/2012/3/e77/>

doi: [10.2196/jmir.2148](https://doi.org/10.2196/jmir.2148)

PMID: [22634568](https://pubmed.ncbi.nlm.nih.gov/22634568/)

©John W Ayers, Benjamin M Althouse, Jon-Patrick Allem, Daniel E Ford, Kurt M Ribisl, Joanna E Cohen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Accuracy of Geographically Targeted Internet Advertisements on Google Adwords for Recruitment in a Randomized Trial

Ray B Jones¹, PhD; Lesley Goldsmith¹, PhD; Christopher J Williams², MD, FRCPsych; Maged N Kamel Boulos¹, MBBCh, PhD, SMIEEE

¹Faculty of Health, Education, and Society, University of Plymouth, Plymouth, United Kingdom

²Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:

Ray B Jones, PhD

Faculty of Health, Education, and Society

University of Plymouth

3 Portland Villas

Plymouth, PL4 8AA

United Kingdom

Phone: 44 1752586532

Fax: 44 1752586532

Email: ray.jones@plymouth.ac.uk

Abstract

Background: Google AdWords are increasingly used to recruit people into research studies and clinical services. They offer the potential to recruit from targeted control areas in cluster randomized controlled trials (RCTs), but little is known about the feasibility of accurately targeting ads by location and comparing with control areas.

Objective: To examine the accuracy and contamination of control areas by a location-targeted online intervention using Google AdWords in a pilot cluster RCT.

Methods: Based on previous use of online cognitive behavioral therapy for depression and population size, we purposively selected 16 of the 121 British postcode areas and randomized them to three intervention and one (do-nothing) control arms. Two intervention arms included use of location-targeted AdWords, and we compared these with the do-nothing control arm. We did not raise the visibility of our research website to normal Web searches. Users who clicked on the ad were directed to our project website, which collected the computer Internet protocol (IP) address, date, and time. Visitors were asked for their postcode area and to complete the Patient Health Questionnaire (depression). They were then offered links to several online depression resources. Google Analytics largely uses IP methods to estimate location, but AdWords uses additional information. We compared locations assessed by (1) Analytics, and (2) as self-identified by users.

Results: Ads were shown 300,523 times with 4207 click-throughs. There were few site visits except through AdWord click-throughs. Both methods of location assessment agreed there was little contamination of control areas. According to Analytics, 69.75% (2617/3752) of participants were in intervention areas, only 0% (8/3752) in control areas, but 30.04% (1127/3752) in other areas. However, according to user-stated postcodes, only 20.7% (463/2237) were in intervention areas, 1% (22/2236) in control areas, but 78.31% (1751/2236) in other areas. Both location assessments suggested most leakage from the intervention arms was to nearby postcode areas. Analytics data differed from postcodes reported by participants. Analysis of a subset of 200/2236 records over 10 days comparing IP-estimated location with stated postcode suggested that Google AdWords targeted correctly in just half the cases. Analytics agreed with our assessment that, overall, one-third were wrongly targeted by AdWords. There appeared little evidence that people who bothered to give their postcode did not answer truthfully.

Conclusions: Although there is likely to be substantial leakage from the targeted areas, if intervention and control areas are a sufficient distance apart, it is feasible to conduct a cluster RCT using online ads to target British postcode areas without significant contamination.

Trial Registration: Clinicaltrials.gov NCT01469689; <http://clinicaltrials.gov/ct2/show/NCT01469689> (Archived by WebCite at <http://www.webcitation.org/681iro5OU>)

(*J Med Internet Res* 2012;14(3):e84) doi:[10.2196/jmir.1991](https://doi.org/10.2196/jmir.1991)

KEYWORDS

Cluster randomized trial; contamination in RCTs; online advertising; depression; MoodGym; Living Life to the Full; Google Analytics; Google Adwords

Introduction

There is evidence that online health interventions can be effective [1], with online cognitive behavioral therapy offering effective treatment for anxiety and depression [2]. Packages such as MoodGYM [3] and Living Life To The Full [4] (LLTTF) are free and available globally. LLTTF is widely used in the United Kingdom and has been demonstrated to be effective, reducing depression for those recruited to use the online package [5], but we previously found a marked geographical variation in uptake. We analyzed distributions of self-reported postcodes of users of LLTTF in Britain over 1 year and found a 15-fold variation that was most likely explained by lack of awareness in some areas (see [Multimedia Appendix 1](#)). A Canadian study found lack of awareness to be the main barrier to effective use of online therapy for depression and other therapies [6].

Depression is a common condition that creates significant workloads for general practice and could be addressed at a population level. While primary care has a major role in health promotion, and some screening activities are run on a population basis from general practice, other public health initiatives have used mass media campaigns [7-9] and online advertising [10], but it is not clear how cost effective online promotion is. Targeted advertising is potentially important for clinical services wanting to contact people in their catchment areas, and in research also to allow geographically matched control arms to be identified.

While randomized controlled trials (RCTs) of online therapies can be carried out by recruitment and random allocation of individuals, the only rigorous way of comparing methods of raising awareness of online therapies is by geolocated cluster RCTs. Studies that look before and after an intervention at global level have no control group that can be adequately described and matched. So studies that simply recruit anyone on the Internet cannot safely match against a defined control area. As a consequence, any increase or decrease in uptake of a therapy could have happened for other reasons. For example, we know that mass media events (eg, a celebrity with a condition, or some other reason for a condition to be in the national news) may have an effect on uptake. Similarly, changes in health services, such as a rollout of measures to improve use of a therapy, or guidelines from a national body may have an impact in uptake nationally over time. Also, access to online therapies is affected by employment and deprivation rates [11]. Matching areas of intervention and control reduces the chances of bias. If we can limit online advertising to one area and compare it with another area where there is no advertising, then we can assume that any difference between the two geographical areas is due to the advertising. We can therefore estimate its cost effectiveness. Subsequently, we can decide whether it is worth using online advertising to raise awareness, or whether other methods would be more cost effective.

Google AdWords (Google Inc, Mountain View, CA, USA) provide the option of purchasing sponsored links that become visible when certain key words are targeted. They can be purchased to be targeted by location. AdWords have been used by others to recruit to studies, for example, to a depression screening site [12], for use of condoms [13], and to a quit-smoking campaign [14], but the cost effectiveness of their use was not assessed. To be able to carry out geolocated cluster RCTs, we need to know whether we can restrict the method of raising awareness to a particular area. If the intervention was via a group of general practitioners or through locally placed advertising, the contamination to other geographical areas is likely to be small. We have therefore run a pilot cluster RCT comparing AdWords targeting depression with local organization website ads and with no intervention, to increase the uptake of online cognitive behavioral therapy, in particular LLTTF [15].

Contamination between intervention and control arms is always of concern in RCTs and often one of the reasons for suggesting cluster rather than individual randomization. But cluster trials also have their limitations [16-21]. On the other hand, interventions, such as population-level advertising, can be randomized only at a cluster rather than an individual level. There seems to have been little examination of contamination between geographical clusters in this sort of trial. We needed to know whether it is possible to target online ads so that there is minimum leakage. Do services such as AdWords target locations as well as claimed? Can we be sure that normal online use of search engines does not corrupt the study? Will leakage be to adjacent areas or, given the increase in use of mobile access and other problems of geolocation, will it be random?

To our knowledge, this is the first study addressing the use of such approaches to target depression. We have assessed contamination between the two interventions that included online ads and control areas and discuss whether geographically targeted online interventions are possible for cluster RCTs.

Methods

Previous registrants on LLTTF gave the first part of their postcode. We stratified postcodes according to their use of LLTTF and population ([Figure 1](#) and [Multimedia Appendix 1](#)) and chose a purposive sample of 16 postcode areas that were similar in population and previous use of LLTTF but were, as far as possible, not adjacent. ([Multimedia Appendix 2](#) shows adjacent postcode areas.) These were randomized to the four arms of the pilot cluster RCT ([Table 1](#), [Figure 2](#)). In the eight areas in arms A and C we ran AdWords. In the eight arms in B and C we aimed to place ads on local organization websites such as local universities, general practices, and local authorities. Arm D was a control arm with no intervention. We had little success in placing ads on local websites, with only three sites (Leeds University, Leeds Carers, and Stronsay Limpet—arm B) agreeing in this period.

Table 1. Study areas, showing eight postcode areas (columns A and C) allocated to AdWords and eight postcode areas (columns B and D) allocated to another intervention and control (total population 7,000,564).

Study area characteristic	A (AdWords)	B (local websites)	C (AdWords and local websites)	D (control)
Total population	1,797,192	1,618,281	1,636,920	1,948,171
Postcode area	Liverpool (L)	Leeds (LS)	London SW (SW)	Nottingham (NG)
Estimated population ^a	843,450	737,343	783,340	1,080,230
Previous use of LLTTF ^b (rate per 100,000 population)	57	68	56	59
Approximate distance (miles) to nearest control (center to center)	35 to Oldham	29 to Oldham	130 to Dudley	NA ^c
Postcode area	Redhill (RH)	Southend (SS)	Kingston (KT)	Oldham (OL)
Estimated population	494,414	493,206	490,104	443,800
Previous use of LLTTF (rate per 100,000 population)	39	46	52	43
Approximate distance (miles) to nearest control (center to center)	120 to Dudley	160 to Dudley	130 to Dudley	NA
Postcode area	Lancaster (LA)	Slough (SL)	Darlington (DL)	Dudley (DY)
Estimated population	325,972	337,631	341,488	397,639
Previous use of LLTTF (rate per 100,000 population)	31	23	38	32
Approximate distance (miles) to nearest control (center to center)	45 to Oldham	110 to Dudley	75 to Oldham	NA
Postcode area	Harrogate (HG)	Kirkwall (KW)	Shetland (ZE)	Hebrides (HS)
Estimated population	133,356	50,101	21,988	26,502
Previous use of LLTTF (rate per 100,000 population)	79	246	155	136
Approximate distance (miles) to nearest control (center to center)	50 to Oldham	135 to Hebrides	235 to Hebrides	NA

^a From Office for National Statistics. National Statistics Postcode Directory. November 2006. Version Notes; 2006. <http://www.statistics.gov.uk/geography/downloads/NSPDVersionNotes.pdf>.

^b Living Life To The Full.

^c Not applicable.

The Google ad (Figure 3) targeted eight postcode areas for two of the arms (A and C in Table 1) using AdWords-customized targeting. Targeting postcode areas (eg, KT) was not an option offered by Google. Options did, however, include targeting a radius of 1 mile or more around a postcode district (eg, KT2) or to hand draw a polygon to enclose the area of interest. Hand drawing a polygon to exactly cover postcode areas was quite difficult, so we used a mix of methods. Four postcode areas (HG, KT, RH, SW) were defined using circles of 1 mile radius for all postcode districts within the postcode area, and four (L, LA, DL, and ZE) were defined by hand-drawn polygons. Multimedia Appendix 3 shows the hand-drawn polygons, gives an assessment of how well the postcode district circles correspond to the postcode area boundary for London SW, and demonstrates the effect of this method in urban and rural areas.

We originally asked AdWords to display the ad for the keyword *depression*. AdWords suggested other similar keyword combinations and we accepted all suggestions (Multimedia

Appendix 4). AdWords gives information on the number of impressions by day, keyword, and location. We ran one simple ad (Figure 3) for all locations with a maximum expenditure of £7.50 per day. AdWords decided when to present the ad. Users searching on terms such as *depression* and *depression help* would, depending on our budget and competing ads, be presented with our ad. Those who clicked on the ad were directed to our research website, which collected the computer Internet protocol (IP) address, date, and time. We specifically did not try to raise the visibility of our website to normal Google, Yahoo, or Bing searches.

Visitors were asked for their postcode area and to complete the Patient Health Questionnaire [22] assessing depression. Users were then offered four links: MoodGYM, LLTTF, NHS Choices information on depression [23], and Samaritans [24] (Multimedia Appendix 5). The order in which the links to MoodGYM and LLTTF were presented was randomized, and similarly the order of links to NHS Choices and Samaritans was

randomized. Data collection methods and AdWords targeting were piloted and refined from April 17 to June 8, 2011. (Multimedia Appendix 6 gives details of some changes made.) We analyzed data from our website and Google Analytics (Google Inc) between June 9 and September 30, 2011 to evaluate the accuracy of targeting of the Google AdWords.

We used location data from Analytics (Table 2) and postcode areas as reported by users (Table 3) to estimate contamination between the intervention and control arms. Analytics is linked to AdWords to the extent that analysis can be restricted to people clicking through from an ad. We used users' IP address as captured on our website to compare with their stated postcode as a comparison (Table 4).

We compared the location of the reported town from Analytics of people clicking on the ad within the 16 sample areas. Analytics has an idiosyncratic view of the geography of London, which seems to be reported as either London, Kensington, Lambeth, or Poplar, and various parts of greater London such as Wembley. It looks as if Google may have divided London into sections (West = Kensington, South = Lambeth, East = Poplar, North included as "London") and then classified unknown places as "generic London." However, despite our repeated attempts to obtain clarification, Google did not answer our requests to clarify this. Others have previously commented on this [25].

Analytics, like many other IP databases, returns a town name. We looked up each town given by Analytics on Google Maps, finding a postcode from the central area of that pin on the map. That postcode was deemed the face value area. We then put the postcode of that location into the Freemap Radius around a

Point tool [26], plotting a 2 mile radius around that postcode. If the place name had a postcode that was within 2 miles of one of our sample areas, we allocated it to that area. Postcodes that were not in our sample were allocated to Other (Table 2).

We used the postcode area, as given by participants on our website questionnaire, to compare how many participants were in the intervention areas and how many in other areas, including our control areas (Table 3). We also calculated the rate of registration on our website for each postcode by dividing the number of registrations by the population of the relevant postcode area. Rates were divided into four groups and mapped, showing intervention and control areas (Figure 2).

For most analyses, we took the postcode area given by the registrant on our website as the reference standard. We had no definitive check on whether they entered their postcode area accurately, but we carried out a more detailed analysis of 200 consecutive users who completed our website questionnaire between September 1 and 10, 2011, comparing the location derived from their IP address and their stated postcode area. We compared the estimated location derived from the computer's IP address with the stated postcode area using three IP location websites: (1) www.whatismyip.com, (2) www.whatismyipaddress.com, and (3) www.maxmind.com.

We examined agreement between the IP location methods and then between one of these (maxmind.com) and the stated postcode area. Combining these approaches, we estimated the impact of contamination on control areas.

The study was approved by the UK National Health Service (NHS) South West 2 Research Ethics Committee (Reference 11/H0203/8; February 2011).

Table 2. Google Analytics data showing total visitors (n = 3752) between June 9 and September 30 (by Cost per Click and other), classified using their town into an arm in the trial or as other area (ie, not in the trial).

Trial arm and intervention	Number of visitors (from Analytics)	
	n	%
A (AdWords)	554	15%
B (Website ads)	29	1%
C (AdWords and website ads)	2034	54.21%
D (Control)	8	0%
Other	1127	30.04%
Total	3752	

Table 3. Website data showing total visitors (n = 2236) between June 9 and September 30 (from all sources), classified using the postcode area they gave into an arm in the trial or as other area (ie, not in the trial).

Trial arm and intervention	Number of visitors (from stated postcode)	
	n	%
A (AdWords)	242	10.8%
B (Website ads)	47	2%
C (AdWords and website ads)	174	7.8%
D (Control)	22	1%
Other	1751	78.31%
Total	2236	

Table 4. Comparison of IP data with stated postcodes for a subsample of 200 visitors who completed our website questionnaire, including postcode area, between September 1 and 10, 2011.

		Stated postcode			
		Targeted	Adjacent	Elsewhere	Total
IP postcode	Targeted	16	7	3	26
	Adjacent	5	26	5	36
	London	10	24	50	84
	United Kingdom	3	4	11	18
	Elsewhere	3	10	23	36
	Total	37	71	92	200

Figure 1. Registration rates by UK postcode area on Living Life To The Full between June 2008 and June 2009, showing four quartiles from darkest (top quartile) to white (lowest quartile) and 16 postcode areas (postcode letters; trial arm) sampled for study and randomized to four arms (A–D).

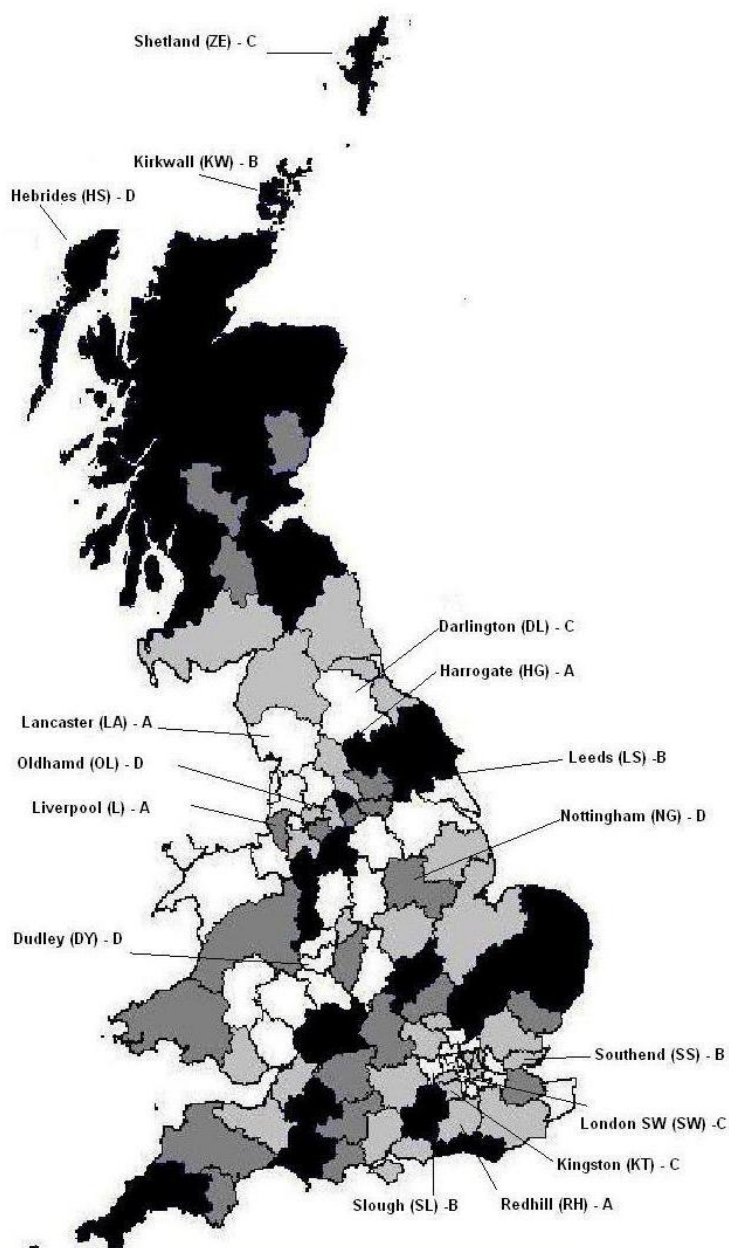


Figure 2. Postcode areas (as stated by participants) showing number of participants (out of 2236) expressed as a rate per million in four groups: darkest, 120+; dark gray, 80 to less than 120; light gray, 40 to less than 80; white, less than 40. Map shows the 16 postcode areas in the study, four in each of arms A, B, C, and D.

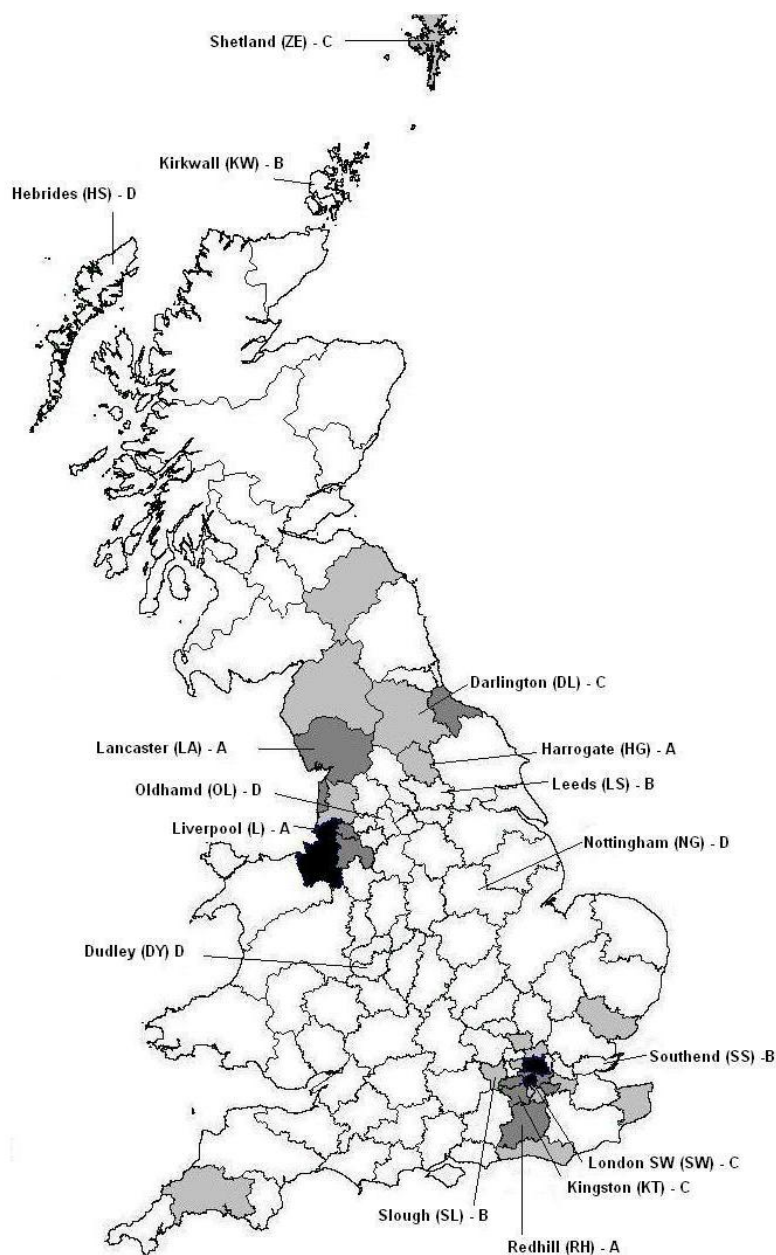
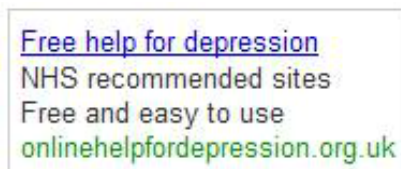


Figure 3. Location-targeted Google ad. NHS = UK National Health Service.



Results

Between June 9 and September 30, 2011, AdWords reported that the ad had been shown 300,523 times and the click-through rate was 1.4%, costing the project £848 at an average cost per click of £0.20. Analytics reported that we had 3929 unique visitors to the website making 4424 visits. Most (96.16%, 4254)

visits were from search engines, of which 4207 (95.07%) were from AdWords. Only 37 visits were from normal searches using Google or Yahoo, 109 were direct traffic (the research team) not included in further analysis, 45 were from one referring site (Leeds University) and another 16 from other sites that were not part of the study. So the site had low visibility and very little access from the wider Web, apart from access by AdWords.

Virtually all (99.21%, 4389) visits were from the United Kingdom, with the few overseas visits likely to be from national Web crawlers. A total of 12% (540) of visits were from mobile devices. The number of clicks on the site gradually increased over the study period (Figure 4).

Tables 2 and 3 show that in the period of study of just under 4 months we had 3752 UK visitors to the website, of whom 2236 (59.59%) completed the website questionnaire giving their postcode. Table 2, based on Analytics estimates of the locations of everyone visiting our website, suggests that there was very little contamination of the control arm in this cluster trial. In the worst-case scenario, only 22/3752 (1%) people who clicked on a Google ad were in the control area. On the other hand, the process was inefficient, as 30% of clicks were wasted on people who were not in the study areas. However, the postcode area stated by participants is our reference standard, and if we look at those who completed our website questionnaire and gave their postcode area (Table 3), we see that only 21% give a postcode in our study area and 1% were in the control area.

Figure 2 shows that most leakage from the intervention arms was to nearby or adjacent postcode areas and that control areas were far enough away not to have much contamination. More leakage seemed to be associated with Liverpool, Lancaster, and Darlington, where we had used polygons to define the areas (Multimedia Appendix 3), and in London SW.

Table 2 shows a major imbalance between the number of participants in arm C and arm A of the trial. Although the arm C intervention was intended to include local website referrals as well as AdWords, we were unable to place many such ads, so arm C was in effect simply AdWords. The target populations of arms A and C were similar, but arm C had nearly four times as many visitors according to Analytics. Table 3, on the other hand, shows that slightly more participants gave their postcode as belonging to arm A than to arm C, but with similar rates per 1000 population (0.13/1000 vs 0.11/1000, respectively).

We know (see above) that 95% of the visitors came to our website as a result of clicking on a Google ad. So, was the

Google ad not well targeted or was the reporting by Analytics inaccurate, or did both contribute? Analytics suggests that 30% of people who clicked on an ad were not in the target areas. But, if we use the postcode areas provided by respondents, 80% were not in target areas (Table 3). However, we can see from Figure 2 that in many cases this was due to leakage to neighboring areas and that contamination of the control areas was only 1%.

To explore further why Analytics and the postcodes stated by participants gave a different picture, we compared the estimated locations from IP addresses. There was good agreement between two of the three location websites (187/200). The IP location given by www.whatismyip.com had little agreement either with the other two databases or with the stated postcode area. We have not used it further. As there was little difference between [whatismyaddress](http://whatismyaddress.com) and [maxmind](http://maxmind.com), we randomly chose [maxmind](http://maxmind.com) to compare against the stated postcodes.

For this subsample of 200/2236 records over 10 days, 19% (37/200) of stated postcodes were in AdWords target areas, which is not dissimilar to that (21%) seen for the whole sample (Table 3). Half of the IP locations were vague (London or United Kingdom). Of the 34 (10 + 24) vague IP locations, 26/34 were for stated postcodes in or adjacent to London SW or Kingston and so could have been in agreement. For all but 1 of the 23 in which both IP location and stated location were off-target, the IP location and self-stated location were different. The 3 where the IP location was on-target and the 5 where the IP location was adjacent but the self-stated postcode was off-target may have been due to the person giving the wrong postcode (eg, their home address rather than current location). The best-case scenario for the accuracy of AdWords targeting is if we assume that the 108 people who gave their postcode as one of the target or adjacent postcodes were on-target or that there was a slight leakage. To this we might add the 8 people thought to be on-target or adjacent by the IP location but whose stated postcode was elsewhere. That is, in total, just over half may have had AdWords targeted correctly.

Figure 4. Google Analytics reporting of the number of clicks on the project website www.onlinehelpfordepression.org.uk every two days from start (April 2011) to mid-October 2011, showing period sampled for this study between June 9 and September 30.



Discussion

This study has shown that a cluster RCT of location-targeted online advertising using AdWords, randomizing at the postcode area level, in Britain is possible without too much contamination. In our pilot cluster RCT, we will ultimately compare the number of registrations on LLTTF in the intervention and control arms,

and the associated costs. However, this intermediate, more detailed analysis of how location targeting works using data from AdWords, Analytics, and our website provides an insight into how to target online advertising by location for researchers considering an RCT using this method.

We had not identified in the literature any rigorous study of raising online awareness in mental health using a cluster RCT

approach with geographical controls. This study has shown that, despite the limitations we encountered in using AdWords, cluster RCTs with geographical controls using British postcode areas are feasible. Contamination between intervention and control arms is not a major problem. In this study we used 12 postcode areas as interventions and 4 as controls, chosen from all 121 British postcode areas, and had less than 1% contamination. Provided the cluster design can have sufficient distance between intervention and control, leakage of location-based online ads should not significantly contaminate the control.

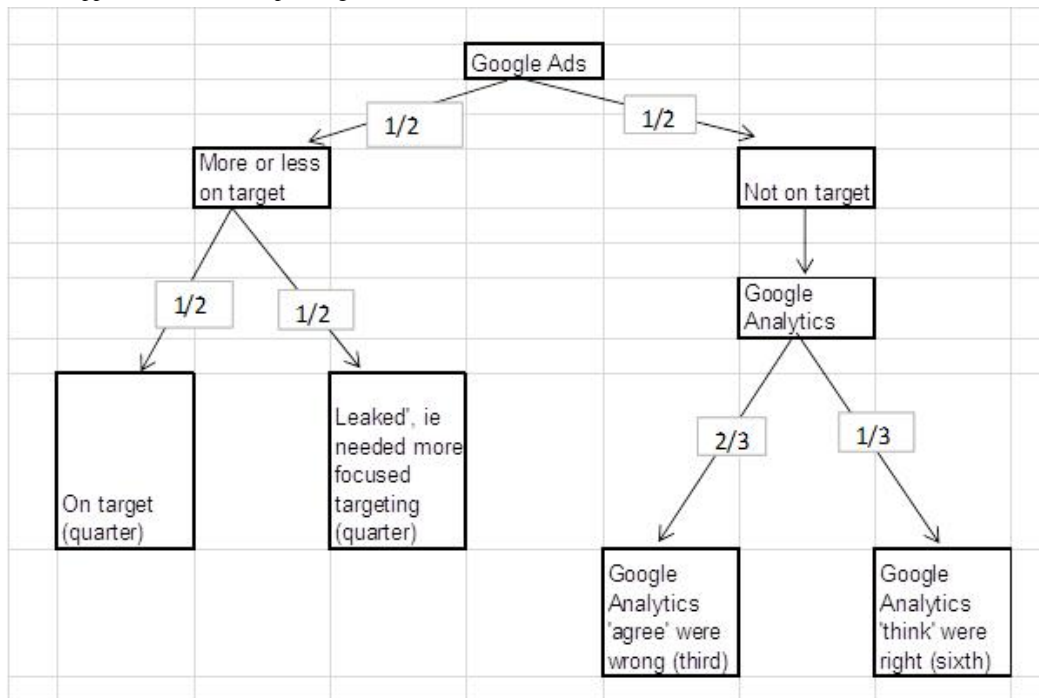
On the other hand, substantial leakage would increase the cost per successful contact. Because of the different denominators, sources of information, and ways of collecting and classifying locations, interpreting Tables 2–4 is quite difficult. At first sight, the numbers may appear contradictory. Figure 5 shows a model of our interpretation of these data. From Table 4, we concluded that about half of the online ads were roughly on-target. We approximated figures from Table 3 to conclude that a quarter were exactly on-target and a quarter had leaked. Overall reporting by Analytics suggested that two-thirds of ads were correctly geographically targeted (Table 2); this implies that Analytics “thinks” that two-thirds to one-half of those off-target are actually right. Thought of in terms of sensitivity and specificity, our campaign was specific, in that no more than 1% (Tables 2) of visitors to our website were from control areas, but was not very sensitive, in that only a quarter (Figure 5) were in target areas.

AdWords currently uses a composite method of IP or user’s address, Web history, and other clues, while Analytics considers only the visitor’s IP address and a lookup database. Bearing this in mind, Table 2 shows that although we asked AdWords to target certain areas, by Google’s own Analytics reporting, it failed in 30% of cases. As Figure 5 shows, some of this failure can be explained by leakage—that is, participants who could have been in the targeted areas if we had used the available methods better—so we cannot blame AdWords for getting it wrong. We estimate that AdWords got the geographical targeting right for about half the people who saw the ads—that is, about half of our participants were actually in the target postcodes

when they clicked on the ad. Of the half whom AdWords did not get on-target, Google (via Analytics) “admitted it” for two-thirds—that is, Analytics recorded them as being in areas that were not targeted by AdWords—whereas Analytics and AdWords were consistent (but wrong) for one-third of those not on-target (ie, one-sixth overall).

Despite the wasted ads, the cost per participant within our intervention areas (approximately £1/person) is still low relative to other methods of raising awareness and is likely to be much more cost effective. For example, in another strand of research in secondary care [27], and more recently in primary care [28], we have been recruiting patients to offer them help in using the Internet. Leafleting methods, particularly in general practice or community settings, cost tens or even hundreds of pounds per person recruited. Of course, if we were not carrying out a cluster trial, AdWords would appear even more cost effective, as raising awareness of online cognitive behavioral therapy is useful wherever people live. The cost of AdWords depends on current competition for the particular keywords being used, so the costs for studies would be different.

The English Department of Health was heavily criticized for spending £2.5 million on AdWords between February 2009 and January 2010 [29]. The Department of Health declined to give further details of how they had spent this money or its cost effectiveness. By judicious design of the website, NHS Choices and other NHS websites should be a high-profile search result. If a Google search returns a website among its first page of normal search results, then advertising using AdWords may be a waste of money. When there is substantial competition for a user’s attention, advertising may be worthwhile. The results of a search and the additional impact of advertising should be explored in more detail before committing to a long-term advertising budget. For example, we examined the probability of finding online cognitive behavioral therapy sites by searching for the term *depression* and have estimated the increased probability as a result of our AdWords campaign (paper submitted). It is possible that the NHS AdWords campaign may have been implemented without prior piloting or evaluation. It may have been cost effective but we do not know.

Figure 5. Approximate model explaining how the data we collected are consistent.

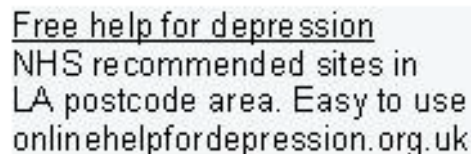
What Have We Learned?

Have we learned helpful points about using AdWords (in its current format) that may be of use to other researchers? Arm C, which included London SW, had (according to Analytics) nearly four times the number of registrants on our website compared with arm A. The population of each was similar, so we must hypothesize that (1) AdWords showed the ad disproportionately (for a given number of people searching on *depression*) more often in London, resulting in more hits, or (2) there were more people in London searching on *depression*, resulting in AdWords displaying it more often, or (3) there was more competition for ads in areas other than London, or (4) more people in London than in other areas clicked through on the ad.

On the other hand, when we examined our website data, we did not see this same imbalance, suggesting two further explanations: (5) that AdWords worked quite well in targeting on the requested areas but Analytics reported very badly, or (6) that many more people from London than from other regions logged on to the site but did not complete the questionnaire. AdWords supplies the number of impressions only by keyword and not by area for a given campaign, and Analytics does not give information on the number of impressions. We do not know

how AdWords decides to display ads if it is given a choice of regions for a given budget.

AdWords provides click-through rates by campaign or by keyword but not by location within a campaign. As we had our AdWords campaign set up initially—as one campaign that included the various locations—we could not analyze click-through rate by location. In retrospect, a better strategy to running one AdWords campaign with multiple postcode areas would be to run separate campaigns for each of the target postcode areas with their own separate budgets allocated proportionately to their populations. This would give more information about why ads were performing better in parts of an intervention arm and would allow for greater control. Furthermore, the postcode area could be included in the ad wording to help in location targeting (Figure 6). After completing data collection for this study, we ran regional ads of this sort for 2 months. Data presented in Multimedia Appendix 7 show that AdWords presented the ad more than twice as often in London SW as in Liverpool per head of population than in other areas. There was no major difference in click-through rates. Running separate campaigns in each area appears to be a better strategy in being able to control a study, but questions remain as to why there should be such a large variation in presentation of ads.

Figure 6. Alternative Google ad for the LA (Lancaster, UK) postcode area.


Free help for depression
NHS recommended sites in
LA postcode area. Easy to use
onlinehelpfordepression.org.uk

Another alternative is to consider excluding London postcodes from the study. However, this would mean that the capital, with its different population and environment compared with many parts of Britain, would not be represented. London distorts the targeting, perhaps because of the likelihood that Internet service providers are based in London or because of the idiosyncratic way that IP location databases and Google interpret the geography of London [25]. Whatever strategy is adopted, the imbalance in the arms points to the need to pilot the study design.

When we started this study, AdWords offered targeting through drawing a radius around a point and a hand-drawn polygon. We used both methods but probably had more leakage from using polygons. Google has (as of October 29, 2011) abandoned the use of hand-drawn polygons [30]. Given our experience, in future (1) we would use only the radius method with the smallest radius possible (1 km) rather than 1 mile in urban areas, (2) in urban areas, we might consider targeting the center of each sample area, leaving a no-man's-land of untargeted areas near the postcode area boundary, and (3) in rural areas, on the other hand, we might use larger-radius circles, as the surface area of postcode districts varies greatly between rural and urban areas (see [Multimedia Appendix 3](#)).

Recommendations for Future Trials

In summary, we would recommend running separate campaigns for different areas with budgets proportional to the target populations, so that the number of presentations of ads and click-through rates can be more easily monitored. We would use radius around a point rather than hand-drawn polygons, and would use the smallest radius (1 km) possible in urban areas. We might avoid including London in a trial and we would put more emphasis on distance between areas included in the study, in particular, to control areas, and less emphasis on trying to match on other characteristics.

Considerations for Improvements in AdWords and Analytics to Support Research

Supporting health research is not Google's main priority, but if we were able to see improvements in Google products, what would be useful for studies of this type? First, we would like

AdWords and Analytics to coordinate better. For example, if we request AdWords to target an area X, we would expect Analytics reports to show that area X was targeted. [Table 2](#) shows that this is currently not the case. Others have commented on the different systems used by AdWords and Google Places [31]. Second, we would like to be able to understand the geography used, particularly of London. Third, we would like more integration between Google Insights and Google AdWords, so that we could understand better whether getting more presentations of an ad in one area was due to the volume of users searching on relevant terms or some function of competing bids.

Limitations

There is of course no guarantee that the ability to target online advertising will always be available in this form. While the increasing use of mobile access to the Internet is encouraging the use of location-specific ads [32], concerns about privacy (eg, [33,34]) and the changing legal environment have resulted in Google allowing people to opt out of its location service [35]. In this study, 12% of visits were from mobiles; it is possible that as this proportion increases, locating users may become more difficult, resulting in more leakage.

AdWords is of course not the only way of local advertising online. This is a complex domain with large financial gains or losses to be made between the big corporations. Ongoing comparisons of the various geotargeted advertising solutions from Microsoft, Facebook, LinkedIn, etc would be worthwhile. For example, Facebook allows geotargeting of ads by radius and other useful criteria such as age, demographics, and interests [36]. Microsoft is now offering radius targeting for Bing and Yahoo with a radius of 5 to 100 miles [37]. Ongoing comparisons are particularly important if we are to ensure that the appropriate population is targeted for a study.

Google has 60%–80% of the search market globally [38]. Facebook is used by 43% of the UK population [39], although it may be underrepresented among, for example, older people or people with few social contacts (who may be the population that most needs to be targeted). Google is starting to compete with Facebook by promoting Google+ [40], where users can

volunteer as much information as they wish about themselves and their current and past locations, and so Google may revise and refine its geotargeting strategy in due course. Each company will continue to develop its methods, sometimes leading to improved, sometimes to reduced, options for researchers.

Cluster trials are stronger—that is, less subject to chance bias—if they include more clusters [21]. Our study, although selection of our clusters took into account population size and previous use of LLTTF, was therefore relatively weak in that each arm included only four postcode areas (clusters). There are two ways we could have included more clusters. First, we could have included more postcode areas. Britain has 121 postcode areas, of which 16 were in our study. Given the shape of Britain, the shape of the postcode areas, the fact that we purposively sampled by quartiles of previous use of LLTTF and population size, and that we had four arms, we already had some difficulty in selecting areas that were not adjacent. However, a simpler trial with two arms and more areas sampled should be possible and would give a stronger design. In a larger country such as the United States it may be easier to choose control arms with no contamination.

The alternative way of increasing the number of clusters is to reduce the target size to a postcode district—that is, to target, for example, KT1, SW4—with controls at the district level. Britain has just over 3000 postcode districts such as this. In our use of AdWords, we defined postcode areas by the sum of postcode districts. Leakage between postcode districts within an area (eg, between KT1 and KT2) was not a problem for our design, in that leakage was into another district in the same geographical area. If we had used single postcode districts scattered over Britain, there would have been more leakage into non-study areas and it would have been difficult to find suitably distant control areas. There is clearly a trade-off between reducing the bias from having a small number of clusters and

having leakage on a greater number of clusters. Analysis of registration on LLTTF by postcode district will allow assessment of intracluster correlation. We think that postcode area is the best size of region for this type of study in the United Kingdom, but having only two arms with more areas in each would be preferable.

We cannot be sure how the findings of this study would translate to other countries, particularly in relation to the complications of including London described above. It seems quite likely that densely populated cities that host many Internet service providers may cause some distortion of the geography used by online advertising.

Other limitations to this study relate to the lack of information from Google or inconsistencies in methods. For example, in trying to assess leakage, we used Google Maps to pin a location for a town and to subsequently draw a radius and assess whether the area of someone using our website might have been targeted by AdWords. But, as we have seen, Google is not consistent between Ads and Analytics, or between Ads and Places, so is quite likely also not to be consistent between Maps and Analytics. However, in the absence of an integrated system, this method did give us an idea of whether we could attribute a case to leakage or to random error.

Conclusions

At least for now, it seems feasible to carry out a cluster RCT of location-targeted online ads in Britain at the postcode area level. Internet recruitment provides unique challenges in understanding the characteristics of participants and which denominators and populations to use [41]. We have highlighted several issues that may help other researchers to use AdWords and other online advertising, but we conclude that it is possible to geographically target Internet ads in a cluster RCT without too much contamination.

Acknowledgments

We would like to thank Ian Mayer of Mirata Ltd for help in extracting data from LLTTF, Robert Stillwell for developing the project website, and Rod Taylor for statistical advice.

The project was funded by a grant from the BUPA Foundation Philip-Poole Wilson Seed Corn Fund.

Ethics committee approval reference: 11-H0203-8.

Authors' Contributions

RJ had the idea for the study, is principal investigator and grant holder, was responsible for day-to-day management of the project, carried out most analyses, and wrote and edited the paper.

LG contributed to the research proposal, is co-grant holder, undertook day-to-day management, did some of the analyses, and helped write and edit the paper.

CW contributed to the research proposal, is co-grant holder, and edited the paper.

MNKB advised on the interpretation of the data, and co-wrote and edited the paper.

Conflicts of Interest

Chris Williams is designer and author of the LLTTF site. LLTTF is run under license by the charity Jubilee Trust.

Multimedia Appendix 1

How the postcode area samples were chosen.

[[PDF File \(Adobe PDF File\), 68KB - jmir_v14i3e84_app1.pdf](#)]

Multimedia Appendix 2

Adjacent areas to the sampled postcode areas.

[[PDF File \(Adobe PDF File\), 43KB - jmir_v14i3e84_app2.pdf](#)]

Multimedia Appendix 3

Further information on boundaries.

[[PDF File \(Adobe PDF File\), 54KB - jmir_v14i3e84_app3.pdf](#)]

Multimedia Appendix 4

Keywords.

[[PDF File \(Adobe PDF File\), 28KB - jmir_v14i3e84_app4.pdf](#)]

Multimedia Appendix 5

Screen shots from research website.

[[PDF File \(Adobe PDF File\), 34KB - jmir_v14i3e84_app5.pdf](#)]

Multimedia Appendix 6

Change made as a result of the pilot period.

[[PDF File \(Adobe PDF File\), 27KB - jmir_v14i3e84_app6.pdf](#)]

Multimedia Appendix 7

Google Adwords budget, clicks, and impressions in two phases of the study.

[[PDF File \(Adobe PDF File\), 38KB - jmir_v14i3e84_app7.pdf](#)]

References

1. Jones R, Goldsmith L. University of Plymouth. 2009 Nov. What is the Evidence for the Benefits and Outcomes for Digital Health Services? Final Report to NHS Choices URL: http://www.plymouth.ac.uk/files/extranet/docs/FoH/NHSChoices_2009.doc [accessed 2012-05-30] [[WebCite Cache ID 683RmVnMa](#)]
2. Spek V, Cuijpers P, Nyklíček 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](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]
3. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ* 2004 Jan 31;328(7434):265. [doi: [10.1136/bmj.37945.566632.EE](https://doi.org/10.1136/bmj.37945.566632.EE)] [Medline: [14742346](https://pubmed.ncbi.nlm.nih.gov/14742346/)]
4. Williams C. Mirata Ltd. 2012 May 31. Living Life To The Full URL: <http://www.lltf.com/> [accessed 2011-10-31] [[WebCite Cache ID 62qfl9MIT](#)]
5. Williams C, Wilson P, Walker A, Wallace I, Morrison J, Whitfield G, et al. An Evaluation of the Effectiveness of Structured Cognitive Behaviour Therapy Self-help Materials Delivered by a Self-help Support Worker Within Primary Care. Report Submitted to the Chief Scientific Officer in Scotland (contract number CZH/4/61). Glasgow: University of Glasgow; 2007.
6. Kakuma R. Utilization of Health Services for Depression and Anxiety in Ontario: An Eleven-Year Comparison of Determinants [thesis]. Montreal: McGill University; 2007.
7. Cram P, Fendrick AM, Inadomi J, Cowen ME, Carpenter D, Vijan S. The impact of a celebrity promotional campaign on the use of colon cancer screening: the Katie Couric effect. *Arch Intern Med* 2003 Jul 14;163(13):1601-1605. [doi: [10.1001/archinte.163.13.1601](https://doi.org/10.1001/archinte.163.13.1601)] [Medline: [12860585](https://pubmed.ncbi.nlm.nih.gov/12860585/)]
8. Flora JA, Maibach EW, Maccoby N. The role of media across four levels of health promotion intervention. *Annu Rev Public Health* 1989;10:181-201. [doi: [10.1146/annurev.pu.10.050189.001145](https://doi.org/10.1146/annurev.pu.10.050189.001145)] [Medline: [2655630](https://pubmed.ncbi.nlm.nih.gov/2655630/)]
9. Howe A, Owen-Smith V, Richardson J. The impact of a television soap opera on the NHS Cervical Screening Programme in the North West of England. *J Public Health Med* 2002 Dec;24(4):299-304 [[FREE Full text](#)] [Medline: [12546207](https://pubmed.ncbi.nlm.nih.gov/12546207/)]
10. Graham AL, Milner P, Saul JE, Pfaff L. Online advertising as a public health and recruitment tool: comparison of different media campaigns to increase demand for smoking cessation interventions. *J Med Internet Res* 2008;10(5):e50 [[FREE Full text](#)] [doi: [10.2196/jmir.1001](https://doi.org/10.2196/jmir.1001)] [Medline: [19073542](https://pubmed.ncbi.nlm.nih.gov/19073542/)]

11. Healthier Scotland. Scottish Executive. 2006 Dec. Delivering for Mental Health URL: <http://www.scotland.gov.uk/Resource/Doc/157157/0042281.pdf> [accessed 2012-01-09] [WebCite Cache ID 64YtHG7OJ]
12. Leykin Y, Muñoz RF, Contreras O. Are consumers of internet health information "cyberchondriacs"? Characteristics of 24,965 users of a depression screening site. *Depress Anxiety* 2012 Jan;29(1):71-77. [doi: [10.1002/da.20848](https://doi.org/10.1002/da.20848)] [Medline: [21681872](https://pubmed.ncbi.nlm.nih.gov/21681872/)]
13. Purdy CH. Using the Internet and social media to promote condom use in Turkey. *Reprod Health Matters* 2011 May;19(37):157-165. [doi: [10.1016/S0968-8080\(11\)37549-0](https://doi.org/10.1016/S0968-8080(11)37549-0)] [Medline: [21555096](https://pubmed.ncbi.nlm.nih.gov/21555096/)]
14. Gordon JS, Akers L, Severson HH, Danaher BG, Boles SM. Successful participant recruitment strategies for an online smokeless tobacco cessation program. *Nicotine Tob Res* 2006 Dec;8 Suppl 1:S35-S41. [Medline: [17491169](https://pubmed.ncbi.nlm.nih.gov/17491169/)]
15. Jones R, Goldsmith L, Williams C. BUPA Foundation. 2010. Raising Awareness of Online Therapies for Depression: Pilot Study URL: <http://www.bupafoundation.co.uk/Seed-Corn-Fund/2010-seed-corn> [accessed 2012-05-30] [WebCite Cache ID 683SDbmCP]
16. Chuang JH, Hripcsak G, Heitjan DF. Design and analysis of controlled trials in naturally clustered environments: implications for medical informatics. *J Am Med Inform Assoc* 2002;9(3):230-238 [FREE Full text] [Medline: [11971884](https://pubmed.ncbi.nlm.nih.gov/11971884/)]
17. Howe A, Keogh-Brown M, Miles S, Bachmann M. Expert consensus on contamination in educational trials elicited by a Delphi exercise. *Med Educ* 2007 Feb;41(2):196-204. [doi: [10.1111/j.1365-2929.2006.02674.x](https://doi.org/10.1111/j.1365-2929.2006.02674.x)] [Medline: [17269954](https://pubmed.ncbi.nlm.nih.gov/17269954/)]
18. Keogh-Brown MR, Bachmann MO, Shepstone L, Hewitt C, Howe A, Ramsay CR, et al. Contamination in trials of educational interventions. *Health Technol Assess* 2007 Oct;11(43):iii, ix-iii,107 [FREE Full text] [Medline: [17935683](https://pubmed.ncbi.nlm.nih.gov/17935683/)]
19. Rhoads CH. The Implications of. *J Educ Behav Stat* 2010;36:76-104. [doi: [10.3102/1076998610379133](https://doi.org/10.3102/1076998610379133)]
20. Richards DA, Lovell K, Gilbody S, Gask L, Torgerson D, Barkham M, et al. Collaborative care for depression in UK primary care: a randomized controlled trial. *Psychol Med* 2008 Feb;38(2):279-287. [doi: [10.1017/S0033291707001365](https://doi.org/10.1017/S0033291707001365)] [Medline: [17803837](https://pubmed.ncbi.nlm.nih.gov/17803837/)]
21. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001 Feb 10;322(7282):355-357. [Medline: [11159665](https://pubmed.ncbi.nlm.nih.gov/11159665/)]
22. Kroenke K, Spitzer RL. The PHQ9: a new depression diagnostic and severity measure. *Psychiatr Ann* 2002;32:509-521.
23. NHS Choices. UK National Health Service. Depression URL: http://www.nhs.uk/conditions/depression/Pages/Introduction.aspx?WT.mc_id=61006. [accessed 2011-10-31] [WebCite Cache ID 62qgIBeRw]
24. Samaritans. 2011. Samaritans URL: <http://www.samaritans.org/> [accessed 2011-10-31] [WebCite Cache ID 62qgVnRjR]
25. Mehew T. SEOptimise. Google Analytics Location Tracking Changes: Where Has London Gone? 2011 Mar 15 URL: <http://www.seoptimise.com/blog/2011/03/google-analytics-location-tracking-changes-where-has-London-gone.html> [accessed 2011-10-31] [WebCite Cache ID 62qggH1xA]
26. Free Map Tools. 2011. Radius From UK Postcode URL: <http://www.freemaptools.com/radius-from-uk-postcode.htm> [accessed 2011-10-31] [WebCite Cache ID 62qguvqLO]
27. Sheaves B, Jones RB, Williamson GR, Chauhan R. Phase 1 pilot study of e-mail support for people with long term conditions using the Internet. *BMC Med Inform Decis Mak* 2011;11:20 [FREE Full text] [doi: [10.1186/1472-6947-11-20](https://doi.org/10.1186/1472-6947-11-20)] [Medline: [21466699](https://pubmed.ncbi.nlm.nih.gov/21466699/)]
28. Jones RB, O'Connor A, Brelsford J, Parsons N, Skirton H. Costs and difficulties of recruiting patients to provide e-health support: pilot study in one primary care trust. *BMC Med Inform Decis Mak* 2012;12(1):25 [FREE Full text] [doi: [10.1186/1472-6947-12-25](https://doi.org/10.1186/1472-6947-12-25)] [Medline: [22458706](https://pubmed.ncbi.nlm.nih.gov/22458706/)]
29. publictechnology.net. 2010 Mar 14. NHS Online Google Ads Come Under Fire URL: <http://www.publictechnology.net/sector/nhs-health/nhs-online-google-ads-come-under-fire> [accessed 2011-10-31] [WebCite Cache ID 62qhBgd4R]
30. Custwin Blogging Online. 2011 Sep 05. Google Adwords Geo Targeting Change URL: <http://www.custwin.co.uk/custwin-blog/2011/09/google-adwords-geo-targeting-change/> [accessed 2011-10-31] [WebCite Cache ID 62qhP8KJB]
31. Altoft P. Blogstorm. 2010 Nov 1. The Google Geolocation Nightmare URL: <http://www.blogstorm.co.uk/the-google-geolocation-nightmare/> [accessed 2012-05-30] [WebCite Cache ID 683TqRVXm]
32. Boulos MN, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. *Biomed Eng Online* 2011;10:24 [FREE Full text] [doi: [10.1186/1475-925X-10-24](https://doi.org/10.1186/1475-925X-10-24)] [Medline: [21466669](https://pubmed.ncbi.nlm.nih.gov/21466669/)]
33. Heimerl JL. Security Week. 2011 Oct 14. A Day in the Life of Privacy URL: <http://www.securityweek.com/day-life-privacy> [accessed 2012-05-30] [WebCite Cache ID 683U9XYm3]
34. Boulos MN. Location-based health information services: a new paradigm in personalised information delivery. *Int J Health Geogr* 2003 Jan 10;2:2 [FREE Full text] [Medline: [12556243](https://pubmed.ncbi.nlm.nih.gov/12556243/)]
35. Sayer P. PCWorld. 2011 Sep 13. Google Lets Wi-Fi Hotspot Owners Opt Out of Location Service URL: http://www.pcworld.com/businesscenter/article/239884/google_lets_wifi_hotspot_owners_opt_out_of_location_service.html [accessed 2011-10-31] [WebCite Cache ID 62qk0Y6W5]
36. Facebook. 2011. Ads and Business Solutions: Targeting Options URL: <http://www.facebook.com/help/?page=863> [WebCite Cache ID 62qkaizj5]
37. Microsoft. 2011. Local Advertising URL: <http://advertising.microsoft.com/small-business/local-advertising> [accessed 2011-10-31] [WebCite Cache ID 62qksVKIz]

38. Netmarketshare. 2011 Sep. Search Engine Market Share URL: <http://marketshare.hitslink.com/search-engine-market-share.aspx?spider=1&qprid=4> [WebCite Cache ID 62ql00Y92]
39. Arthur C. The Guardian. 2010 Jul 22. You Want an Interactive Map of Where Facebook is Used? Happy to Help URL: <http://www.guardian.co.uk/technology/blog/2010/jul/22/facebook-countries-population-use> [accessed 2011-10-31] [WebCite Cache ID 62qlGbORF]
40. Google. 2011. Google+ URL: <https://plus.google.com/> [WebCite Cache ID 62qlQO5kM]
41. Graham AL, Bock BC, Cobb NK, Niaura R, Abrams DB. Characteristics of smokers reached and recruited to an internet smoking cessation trial: a case of denominators. *Nicotine Tob Res* 2006 Dec;8 Suppl 1:S43-S48. [Medline: [17491170](#)]

Abbreviations

IP: Internet protocol

LLTTF: Living Life To The Full

NHS: UK National Health Service

RCT: randomized controlled trial

Edited by T Houston; submitted 08.11.11; peer-reviewed by R Hsiung, W van Ballegooijen; comments to author 26.12.11; revised version received 09.01.12; accepted 13.04.12; published 20.06.12.

Please cite as:

Jones RB, Goldsmith L, Williams CJ, Kamel Boulos MN

Accuracy of Geographically Targeted Internet Advertisements on Google Adwords for Recruitment in a Randomized Trial

J Med Internet Res 2012;14(3):e84

URL: <http://www.jmir.org/2012/3/e84/>

doi: [10.2196/jmir.1991](https://doi.org/10.2196/jmir.1991)

PMID: [22718043](https://pubmed.ncbi.nlm.nih.gov/22718043/)

©Ray B Jones, Lesley Goldsmith, Christopher J Williams, Maged N Kamel Boulos. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 20.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Short Paper

Clinicians' Perspectives on a Web-Based System for Routine Outcome Monitoring in Old-Age Psychiatry in the Netherlands

Marjolein A Veerbeek¹, MSc; Richard C Oude Voshaar², PhD; Anne Margriet Pot^{1,3}, PhD

¹Netherlands Institute of Mental Health and Addiction, Utrecht, Netherlands

²UMC Groningen, University Center of Psychiatry, Groningen, Netherlands

³Department of Clinical Psychology, VU University, Amsterdam, Netherlands

Corresponding Author:

Marjolein A Veerbeek, MSc

Netherlands Institute of Mental Health and Addiction

PO Box 725

Utrecht, 3500 AS

Netherlands

Phone: 31 30 2959313

Fax: 31 30 2971111

Email: mveerbeek@trimbos.nl

Abstract

Background: In health care, the use of physical parameters to monitor physical disease progress is common. In mental health care, the periodic measurement of a client's functioning during treatment, or routine outcome monitoring, has recently become important. Online delivery of questionnaires has the potential to reduce clinicians' resistance to the implementation of routine outcome monitoring. Online delivery enables clinicians to receive results on a questionnaire in a graphic directly after data entry. This gives them insight into the progress of a client at a single glance.

Objective: To explore clinicians' perspectives on a routine outcome monitoring procedure where questionnaires and feedback on scores were delivered online. Questionnaires could also be filled out on paper and then entered into the online system by a research assistant.

Methods: In 2009 we sent an online survey, consisting of five yes-or-no questions and six open-ended questions, to all clinicians in the 14 mental health care organizations working with the routine outcome monitoring system in the Netherlands. Of the 172 clinicians contacted, 80 (47%) opened the link and 70 of these 80 (88%) clinicians completed the survey.

Results: Clinicians seldom used the graphical feedback from the Web-based system, which indicates that direct feedback on scores did not enhance the implementation of routine outcome monitoring. Integration into the electronic patient record and more training on interpretation and implementation of feedback in daily practice were seen as the primary points for further improvement. It was mainly the availability of a research assistant that made the routine outcome monitoring procedure feasible.

Conclusions: Without a research assistant and training in the interpretation of outcomes, software programs alone cannot ensure effective implementation of monitoring activities in everyday practice.

(*J Med Internet Res* 2012;14(3):e76) doi:[10.2196/jmir.1937](https://doi.org/10.2196/jmir.1937)

KEYWORDS

Quality of care; health care; routine outcome monitoring; Web-based monitoring; clinicians; feedback

Introduction

Regular monitoring of clients with chronic physical diseases using specific outcome measures—for example glucose levels in diabetes or blood pressure in patients with heart failure—is common in health care [1-3]. In mental health care, the periodic measurement of clients' functioning during treatment, or routine outcome monitoring, has only recently become important [4-7].

By providing regular information to both the professional and the client on the course and severity of symptoms during treatment, routine outcome monitoring is assumed to improve informed decision making and therefore quality of care. Well-timed information on the severity and course of symptoms during treatment is an integral part of routine outcome monitoring [8,9]. Increased workload due to systematic data collection, such as filling out questionnaires multiple times, can

easily lead to resistance by clinicians [10]. Therefore, the additional burden of data collection should balance the utility of these data from the clinician's perspective.

Online delivery of questionnaires may facilitate routine outcome monitoring by (1) reducing the logistic burden, (2) being able to provide feedback directly to clinicians, which is an extension of the paper-and-pencil method, and (3) reducing the amount of missing data, because one may not proceed without answering all questions [6,10,11]. However, previous research has suggested that attention should be paid to cross-validation between electronic and paper-and-pencil versions of the questionnaires, as psychometric properties may change [12-14] and Internet-based surveys do not automatically translate into higher response rates than with paper-based versions [15-19].

To stimulate the use of routine outcome monitoring in old-age psychiatry in the Netherlands, we introduced the Mental Health Care Monitor Older Adults (MEMO) system. This Web-based system was designed to deliver the benefits mentioned above and contains Web-based versions of questionnaires, automatically calculated sum scores, graphics presenting the symptom course over time, and reminders to clinicians of when follow-up assessments are needed. Depending on their own individual preferences, clinicians could also fill out the questionnaires on paper and a research assistant would then enter the data into the Web-based system. The present study

explored the acceptability of using a Web-based routine outcome monitoring system in old-age psychiatry from the perspective of the clinician.

Methods

Description of MEMO

MEMO was started in 2008 for a period of 5 years. During information meetings, managers received extensive information about the procedure of MEMO. After these meetings, 14 of the 41 mental health care organizations throughout the Netherlands applied to participate in MEMO. All clients referred by their general practitioner to these 14 organizations and who progressed to treatment after intake at an outpatient clinic for old-age psychiatry were monitored. The only exclusion criterion was a primary diagnosis of a cognitive disorder. Data were collected at fixed intervals and included diagnostic variables, measures of mental and social functioning, treatment characteristics, and client satisfaction (see Table 1 [20-23]).

Before the introduction of MEMO, outcomes were not measured in such a structured and uniform way in old-age psychiatry in the Netherlands. To ensure high-quality data collection, funding was provided to organizations to employ a research assistant. We also provided annual training sessions for the use of the Health of the Nation Outcome Scales for older adults (HoNOS 65+) and the Web-based system (see below).

Table 1. Measurement overview of Mental Health Care Monitor Older Adults (MEMO).

Measure	Rater	Time point			
		Intake	4 months	8 months	12 months/end of treatment
Demographic variables	MHP ^a	Yes	No	No	No
DSM-IV-TR ^b classification	MHP	Yes	No	No	No
HoNOS 65+ ^c (mental and social functioning)	MHP	Yes	Yes	Yes	Yes
GDS-15 ^d (severity of depressive symptoms) ^e	Client	Yes	Yes	Yes	Yes
Type of treatment ^f	MHP	No	Yes	Yes	Yes
Life events	MHP	No	No	No	Yes
MHCCT ^g (client satisfaction)	Client	No	No	No	Yes

^a Mental health care professional.

^b Diagnosis according to the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, text revision [20].

^c Health of the Nation Outcome Scales for older adults [21].

^d Geriatric Depression Scale 15-item version [22].

^e To be filled out only by clients with a primary diagnosis of a mood disorder.

^f Categories are based on the Dutch reimbursement categorization: psychological therapy, supportive counseling, activating techniques, psychopharmaceutical treatment, relation or system therapy, and electroconvulsive therapy.

^g Mental Health Care Client Thermometer [23].

As clients neither were subjected to interventions nor had to obey behavioral rules for MEMO, the medical ethics committee Medisch-ethische Toetsingscommissie instellingen Geestelijke Gezondheidszorg decided that the Dutch Medical Research Involving Human Subjects Act was not applicable and the study did not require ethics approval. Nevertheless, all clients were

informed about the scientific purposes of MEMO, and clients who did not agree to the use of their anonymous data were excluded.

Web-Based Data Collection

NetQuestionnaires Netherlands (Utrecht, the Netherlands; www.netq.nl) developed our Web-based system for both data

acquisition and processing (sum scores and graphical presentation of symptom course).

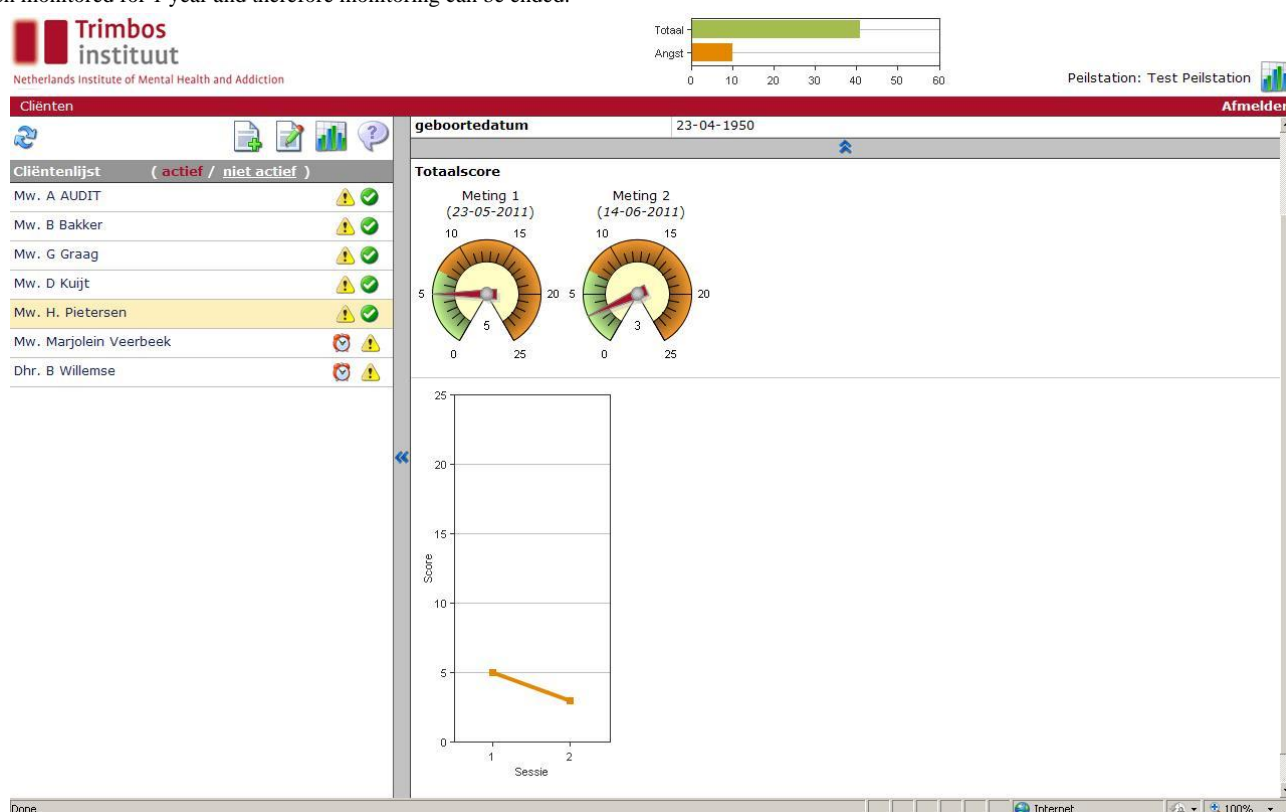
To safeguard the data collected, we observed the Data Protection Act and ensured confidentiality through the use of an HTTPS connection, which secured the communication on the Internet, as well as a login procedure and a separate environment for each organization, which prevented the exchange of client data between organizations [24].

After a client was added to the system, based on the diagnosis, the questionnaires that had to be filled out at baseline were shown and could be filled out immediately online. During

follow-up, the system tracked each client and signaled every 4 months, up to 12 months, by email alerts to both clinicians and research assistants when new questionnaires had to be completed. Immediately after data entry, the system generated a graphic visualizing the progress of the client and thereby giving feedback to the clinician. For example, the feedback in Figure 1 shows that the client had a score of 5 at intake and improved after 4 months to a score of 3.

For clinicians who preferred to use paper-and-pencil forms, feedback was delayed until the research assistant had entered the data into the system. The participating organizations asked clients to fill out their questionnaires on paper.

Figure 1. Example of the graphical feedback generated by the Web-based system to show clinicians the progress of their client during treatment. Clarification of Dutch language terms: *Totaalscore* = sum score (at a particular questionnaire); *Meting 1* = sum score at intake; *Meting 2* = sum score at the second moment in time; *Score* (y-axis) = sum score; *Sessie* (x-axis) = session (1 = intake, 2 = second moment in time). Clarification of symbols: exclamation mark = questionnaires have to be filled out for this particular client; check mark = the client has ended treatment; alarm clock = client has been monitored for 1 year and therefore monitoring can be ended.



Evaluation Questionnaire

Measurement

To evaluate the MEMO procedure, we emailed an online questionnaire to all 172 clinicians who were authorized to use the Web-based system after 1 year of data collection. They were asked to help improve the procedure of MEMO by filling out the questionnaire, which would take 10 minutes at most. Responses would not be communicated to their organization or manager. To avoid duplicate entries the software did not allow clinicians to open the questionnaire more than once. In addition to general information (organization, sex, age, profession, and work experience), the questionnaire consisted of 5 items that could be answered with either yes or no: (1) Do you think the MEMO procedure is feasible? (2) Do you use the graphical

feedback from the system? If yes, the following questions were displayed: (3) Do you think the graphical feedback is relevant to your work? (4) Based on the graphical feedback, did you change the treatment plan for a specific client? (5) Do you think this type of feedback should be continued? Each question was accompanied by an open-ended question that gave them an opportunity to explain their answer. Next, clinicians were asked to grade the graphical feedback (on a scale of 1 to 10). The final question for all clinicians was "Do you have any suggestions for data collection and giving feedback in mental health care for older adults?" The questions were spread over 4 pages: (1) general information, (2) feasibility questions, (3) questions on the graphical feedback, (4) suggestions for data collection. Clinicians were able to review and change their answers in the course of completing the questionnaire. Responses were automatically captured in a database. To improve the response

rate, (1) contrary to the yes-or-no questions, the open-ended questions were not mandatory to prevent premature discontinuation, (2) clinicians received a reminder after 1 month, and (3) a gift voucher of €25 was raffled off among the respondents in each organization.

Analysis

Qualitative data analysis was performed on the answers to the open-ended questions. The contents of the answers were labeled (by MV). These labels were data driven. When coding was complete, the frequencies of the codes were analyzed.

Results

Of the 172 clinicians, 80 (47%) opened the link to the evaluation questionnaire; 70 of these 80 (88%) clinicians completed the questionnaire ($n = 41$, 59% female; mean age 47.5, range 27–64

years; mean work experience 22, range 1–44 years). Participating clinicians were nurses ($n = 38$, 54%), psychologists ($n = 24$, 34%), psychiatrists or geriatricians ($n = 5$, 7%), and social workers ($n = 3$, 4%). Characteristics of nonresponders are unknown.

Although 54 of the 70 (77%) clinicians considered the procedures of MEMO to be feasible, only 11 of 70 (16%) actually used the graphical feedback (Table 2). The 11 clinicians who used the feedback considered that it was relevant to everyday practice and that it should be continued. They rated the quality of the feedback with a score of 7 (range 5–8). Clinicians who did not use the feedback gave the following reasons: no time ($n = 8$), forgot how to use the feedback ($n = 5$), did not understand graphical feedback ($n = 3$), and had recently joined the team and therefore had not yet had the opportunity to measure clients multiple times ($n = 4$).

Table 2. Clinicians' answers to the yes-or-no questions on the feasibility of Mental Health Care Monitor Older Adults (MEMO) and the use of feedback from the Web-based system.

Question	n	Positive answer (yes)	
		n	%
1 Procedure MEMO feasible	70	54	77%
2 Use of graphical feedback	70	11	16%
3 Graphical feedback relevant	11	10	91%
4 Change of treatment plan	11	4	36%
5 Continuation of graphical feedback	11	11	100%

Of 70 clinicians, 21 (30%) gave further suggestions, and 8 of these 21 (38%) suggested that the Web-based system could be improved by incorporating the system within the electronic patient record, through which alerts and graphics are seen directly in the record. The additional login procedure besides the electronic patient record was thought to be the main barrier to the use of the Web-based system. Instead, they filled out the questionnaires on paper. Of the 21 clinicians, 6 (29%) suggested that data collection be extended to other instruments beyond the HoNOS 65+ to obtain more detailed information. Of the 21 clinicians, 5 (24%) suggested that more attention should be paid to the implementation of routine outcome monitoring in everyday practice. Remarkably, only two comments about increased workload were given.

Discussion

Although the majority of clinicians considered the procedure of MEMO to be feasible, only a few actually used the graphical feedback from the Web-based system. The most common suggestion given to enhance the use of the feedback was to incorporate the Web-based system into the electronic patient record, so that the scores and graphics of their clients would be more easily accessible. This means that previous findings that software programs support direct feedback and monitor activities in clinical practice [6] hold true only if these programs are integrated into the electronic patient record.

As we did not integrate our Web-based system into the electronic patient record, it is likely that the availability of a

research assistant made our routine outcome monitoring procedures feasible. This is in line with the findings of a previous study on younger adults, which showed that routine outcome monitoring is highly feasible when supported by trained assistants [7]. However, in that project, the routine outcome monitoring assistants carried out all measurements and reported to the clinicians by letter. In our study, clinicians performed measurements themselves and the research assistants did not report results. The research assistants only entered data into the Web-based system to enable clinicians to see the graphical feedback. Recent research showed that when female cancer survivors were given a choice between filling out a questionnaire online and filling out a paper version of the questionnaire, they preferred to fill out the paper version. When they were not given the choice, their response rates were similar for both the Web-based and paper-based questionnaire [25]. This provides further evidence that it was the availability of a research assistant that made our routine outcome monitoring procedure feasible, since clinicians had a choice and probably preferred to fill out the questionnaires on paper. However, the added value of our research assistants was probably much greater than simple data entry for the following reasons: (1) they had an overview of the entire project, and (2) they were easily accessible to clinicians for support using the Web-based system. Finally, they reminded clinicians to follow up measurements before passing the deadline. Therefore, even when clinicians did not have the opportunity to fill out questionnaires on paper, the support of a research assistant was critical to the implementation of the routine outcome monitoring and is probably the reason that in

our study only two comments on the increased workload were made. Previously reported barriers to routine outcome monitoring implementation from the clinician's perspective included (1) the opinion that the expert knows best, (2) outcome measurement undermining clinical expertise and violating the privacy of the therapy dyad, (3) increased workload, and (4) fear of evaluation of their practice [6,7,10,26,27]. None of these reasons were mentioned in our survey. Instead, a lack of knowledge of how to implement outcome measures in day-to-day clinical practice appeared to be another barrier. Future research should investigate whether more training enhances the use of direct feedback in clinical practice. After successful implementation of MEMO, results will provide a proxy of quality of care. Additional measures of routine outcome monitoring will be needed to indicate quality of care in old-age psychiatry.

Limitations

First, the overall response rate (70/172, 41%) is rather low. Nevertheless, most nonresponders (92/172, 54%) did not open the link and therefore were not aware of the content of the survey. As 88% (70/80) of the clinicians who opened the link also filled out the questionnaire, it seems unlikely that the

response rate is related to the content of the questionnaire. Second, as the explanatory items were not mandatory, some clinicians may have too easily skipped some of these questions. Therefore, reasons for not using the feedback or suggestions to improve the procedure of MEMO may have been missed. Third, open-ended questions were labeled by one person. Since the questions were structured and the clinicians' answers were brief, it is likely that having a second person coding would have led to the same results. Fourth, organizations voluntarily applied for participation in MEMO, which could imply that participating clinicians were more in favor of using a Web-based system than were clinicians in organizations that did not participate. However, this seems unlikely to have been the case, because it was the managers who made the decision to participate in MEMO and not the clinicians themselves.

Conclusion

The availability of a research assistant seems to enhance the acceptability of routine outcome monitoring to clinicians. Integration of routine outcome monitoring software into the electronic patient record and more training on how to implement direct feedback in daily practice would make the Web-based system easier to use for clinicians.

Acknowledgments

This study was financed by the Netherlands Ministry of Health, Welfare and Sport. We are grateful for this financial assistance in enabling MEMO.

Authors' Contributions

MV drafted the manuscript and helped to design the study. ROV helped to draft the manuscript and supervised the study. AMP designed and supervised the study and the present manuscript.

Conflicts of Interest

None declared.

References

1. Wu RC, Delgado D, Costigan J, Maciver J, Ross H. Pilot study of an Internet patient-physician communication tool for heart failure disease management. *J Med Internet Res* 2005 Feb;7(1):e8 [FREE Full text] [doi: [10.2196/jmir.7.1.e8](https://doi.org/10.2196/jmir.7.1.e8)] [Medline: [15829480](https://pubmed.ncbi.nlm.nih.gov/15829480/)]
2. Mazzi CP, Kidd M. A framework for the evaluation of Internet-based diabetes management. *J Med Internet Res* 2002 Feb;4(1):e1 [FREE Full text] [doi: [10.2196/jmir.4.1.e1](https://doi.org/10.2196/jmir.4.1.e1)] [Medline: [11956033](https://pubmed.ncbi.nlm.nih.gov/11956033/)]
3. Gillespie G. Deploying an I.T. cure for chronic diseases. *Health Data Manag* 2000 Jul;8(7):68-70, 72, 74. [Medline: [11067308](https://pubmed.ncbi.nlm.nih.gov/11067308/)]
4. van der Krieke L, Emerencia AC, Aiello M, Sytema S. Usability evaluation of a web-based support system for people with a schizophrenia diagnosis. *J Med Internet Res* 2012 Feb;14(1):e24 [FREE Full text] [doi: [10.2196/jmir.1921](https://doi.org/10.2196/jmir.1921)] [Medline: [22311883](https://pubmed.ncbi.nlm.nih.gov/22311883/)]
5. Gilbody SM, House AO, Sheldon TA. Outcome measures and needs assessment tools for schizophrenia and related disorders. *Cochrane Database Syst Rev* 2003(1):CD003081. [doi: [10.1002/14651858.CD003081](https://doi.org/10.1002/14651858.CD003081)] [Medline: [12535453](https://pubmed.ncbi.nlm.nih.gov/12535453/)]
6. Lambert MJ, Harmon C, Slade K, Whipple JL, Hawkins EJ. Providing feedback to psychotherapists on their patients' progress: clinical results and practice suggestions. *J Clin Psychol* 2005 Feb;61(2):165-174. [doi: [10.1002/jclp.20113](https://doi.org/10.1002/jclp.20113)] [Medline: [15609358](https://pubmed.ncbi.nlm.nih.gov/15609358/)]
7. de Beurs E, den Hollander-Gijsman ME, van Rood YR, van der Wee NJ, Giltay EJ, van Noorden MS, et al. Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clin Psychol Psychother* 2011 Feb;18(1):1-12. [doi: [10.1002/cpp.696](https://doi.org/10.1002/cpp.696)] [Medline: [20238371](https://pubmed.ncbi.nlm.nih.gov/20238371/)]
8. Zwanepol F, de Groot W. [Implementation of routine outcome monitoring: experiences with clients with mood and anxiety disorders]. *Maandblad Geestelijke Volksgezondheid* 2008 Feb;63:118-129.

9. Srebnik D, Hendryx M, Stevenson J, Caverly S, Dyck DG, Cauce AM. Development of outcome indicators for monitoring the quality of public mental health care. *Psychiatr Serv* 1997 Jul;48(7):903-909. [Medline: [9219298](#)]
10. Madan A, Borckardt JJ, Connell A, Book SB, Campbell S, Gwynette MF, et al. Routine assessment of patient-reported outcomes in behavioral health: room for improvement. *Qual Manag Health Care* 2010 Mar;19(1):70-81. [doi: [10.1097/QMH.0b013e3181ccbc53](#)] [Medline: [20042935](#)]
11. de Nooijer J, de Vries NK. Monitoring health risk behavior of Dutch adolescents and the development of health promoting policies and activities: the E-MOVO project. *Health Promot Int* 2007 Mar;22(1):5-10 [FREE Full text] [doi: [10.1093/heapro/dal036](#)] [Medline: [17015406](#)]
12. Shervin N, Dorrwachter J, Bragdon CR, Shervin D, Zurakowski D, Malchau H. Comparison of paper and computer-based questionnaire modes for measuring health outcomes in patients undergoing total hip arthroplasty. *J Bone Joint Surg Am* 2011 Feb 2;93(3):285-293. [doi: [10.2106/JBJS.I.01104](#)] [Medline: [21266642](#)]
13. Touvier M, Méjean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010 May;25(5):287-296. [doi: [10.1007/s10654-010-9433-9](#)] [Medline: [20191377](#)]
14. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. *J Med Internet Res* 2011 Dec;13(4):e109 [FREE Full text] [doi: [10.2196/jmir.1593](#)] [Medline: [22155721](#)]
15. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison (2). *J Med Internet Res* 2004 Oct 29;6(4):e39 [FREE Full text] [doi: [10.2196/jmir.6.4.e39](#)] [Medline: [15631963](#)]
16. McMahon SR, Iwamoto M, Massoudi MS, Yusuf HR, Stevenson JM, David F, et al. Comparison of e-mail, fax, and postal surveys of pediatricians. *Pediatrics* 2003 Apr;111(4 Pt 1):e299-e303. [Medline: [12671142](#)]
17. Kim HL, Gerber GS, Patel RV, Hollowell CM, Bales GT. Practice patterns in the treatment of female urinary incontinence: a postal and internet survey. *Urology* 2001 Jan;57(1):45-48. [Medline: [11164141](#)]
18. Harewood GC, Yacavone RF, Locke GR, Wiersema MJ. Prospective comparison of endoscopy patient satisfaction surveys: e-mail versus standard mail versus telephone. *Am J Gastroenterol* 2001 Dec;96(12):3312-3317. [doi: [10.1111/j.1572-0241.2001.05331.x](#)] [Medline: [11774942](#)]
19. Raziano DB, Jayadevappa R, Valenzula D, Weiner M, Lavizzo-Mourey R. E-mail versus conventional postal mail survey of geriatric chiefs. *Gerontologist* 2001 Dec;41(6):799-804. [Medline: [11723348](#)]
20. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders:DSM-IV-TR*. 4th edition, text revision. Washington, DC: American Psychiatric Association; 2000.
21. Burns A, Beevor A, Lelliott P, Wing J, Blakey A, Orrell M, et al. Health of the Nation Outcome Scales for elderly people (HoNOS 65+). *Br J Psychiatry* 1999 May;174:424-427. [Medline: [10616609](#)]
22. Sheikh RL, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol* 1986;5(1-2):165-173.
23. Kok I, Mulder E. [Client Satisfaction in Mental Health Care: Manual for Different Client Thermometers in Use]. Utrecht: Trimbos-instituut; 2005.
24. Kelly G, McKenzie B. Security, privacy, and confidentiality issues on the Internet. *J Med Internet Res* 2002;4(2):e12 [FREE Full text] [doi: [10.2196/jmir.4.2.e12](#)] [Medline: [12554559](#)]
25. van den Berg MH, Overbeek A, van der Pal HJ, Versluys AB, Bresters D, van Leeuwen FE, et al. Using web-based and paper-based questionnaires for collecting data on fertility issues among female childhood cancer survivors: differences in response characteristics. *J Med Internet Res* 2011;13(3):e76 [FREE Full text] [doi: [10.2196/jmir.1707](#)] [Medline: [21955527](#)]
26. McGrath BM, Tempier RP. Implementing quality management in psychiatry: from theory to practice--shifting focus from process to outcome. *Can J Psychiatry* 2003 Aug;48(7):467-474. [Medline: [12971017](#)]
27. Robertson L, Smith M, Castle D, Tannenbaum D. Using the Internet to enhance the treatment of depression. *Australas Psychiatry* 2006 Dec;14(4):413-417. [doi: [10.1111/j.1440-1665.2006.02315.x](#)] [Medline: [17116083](#)]

Abbreviations

HoNOS 65+: Health of the Nation Outcome Scales for older adults

MEMO: Mental Health Care Monitor Older Adults

Edited by R Crutzen; submitted 14.09.11; peer-reviewed by H Veerbeek, M Jordan-Marsh, S Langrial; comments to author 18.10.11; revised version received 05.12.11; accepted 16.03.12; published 30.05.12.

Please cite as:

Veerbeek MA, Oude Voshaar RC, Pot AM

Clinicians' Perspectives on a Web-Based System for Routine Outcome Monitoring in Old-Age Psychiatry in the Netherlands

J Med Internet Res 2012;14(3):e76

URL: <http://www.jmir.org/2012/3/e76/>

doi: [10.2196/jmir.1937](https://doi.org/10.2196/jmir.1937)

PMID: [22647771](https://pubmed.ncbi.nlm.nih.gov/22647771/)

©Marjolein A Veerbeek, Richard C Oude Voshaar, Anne Margriet Pot. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 30.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Tutorial

Using Crowdsourcing Technology for Testing Multilingual Public Health Promotion Materials

Anne M Turner^{1,2}, MPH, MLIS, MD; Katrin Kirchhoff³, PhD; Daniel Capurro^{2,4}, MD

¹Department of Health Services, University of Washington, Seattle, WA, United States

²Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

³Department of Electrical Engineering, University of Washington, Seattle, WA, United States

⁴Department of Internal Medicine, Pontificia Universidad Catolica de Chile, Santiago, Chile

Corresponding Author:

Anne M Turner, MPH, MLIS, MD

Department of Health Services

University of Washington

1107 NE 45th Street

Suite 400

Seattle, WA, 98105

United States

Phone: 1 2066851130

Fax: 1 2066169415

Email: amturner@uw.edu

Abstract

Background: Effective communication of public health messages is a key strategy for health promotion by public health agencies. Creating effective health promotion materials requires careful message design and feedback from representatives of target populations. This is particularly true when the target audiences are hard to reach as limited English proficiency groups. Traditional methods of soliciting feedback—such as focus groups and convenience sample interviews—are expensive and time consuming. As a result, adequate feedback from target populations is often insufficient due to the time and resource constraints characteristic to public health.

Objective: To describe a pilot study investigating the use of crowdsourcing technology as a method to gather rapid and relevant feedback on the design of health promotion messages for oral health. Our goal was to better describe the demographics of participants responding to a crowdsourcing survey and to test whether crowdsourcing could be used to gather feedback from English-speaking and Spanish-speaking participants in a short period of time and at relatively low costs.

Methods: We developed health promotion materials on pediatric dental health issues in four different formats and in two languages (English and Spanish). We then designed an online survey to elicit feedback on format preferences and made it available in both languages via the Amazon Mechanical Turk crowdsourcing platform.

Results: We surveyed 236 native English-speaking and 163 native Spanish-speaking participants in less than 12 days, at a cost of US \$374. Overall, Spanish-speaking participants originated from a wider distribution of countries than the overall Latino population in the United States. Most participants were in the 18- to 29-year age range and had some college or graduate education. Participants provided valuable input for the health promotion material design.

Conclusions: Our results indicate that crowdsourcing can be an effective method for recruiting and gaining feedback from English-speaking and Spanish-speaking people. Compared with traditional methods, crowdsourcing has the potential to reach more diverse populations than convenience sampling, while substantially reducing the time and cost of gathering participant feedback. More widespread adoption of this method could streamline the development of effective health promotion materials in multiple languages.

(*J Med Internet Res* 2012;14(3):e79) doi:[10.2196/jmir.2063](https://doi.org/10.2196/jmir.2063)

KEYWORDS

Crowdsourcing; health promotion; public health informatics; limited English proficiency populations

Introduction

Effective communication of health messages to a wide range of populations is a key public health strategy for preventing disease. Unfortunately, the vast majority of good-quality health information materials—such as websites, flyers, and patient handouts—in the United States are available only in English. Communication efforts frequently do not reach diverse populations because of language barriers. Design of effective health promotion materials for linguistically diverse populations requires careful development of key messages, as well as evaluation and feedback from target communities. Most methods that require recruitment of participants to develop and test multilingual health messages are both costly and time consuming [1]. We report here on the use of crowdsourcing to gather quick feedback on health promotion materials from English-speaking and Spanish-speaking populations in an inexpensive and efficient manner. The specific advantages and challenges of using this technique for health communications research are explored.

Background and Motivation

Health promotion materials—either in print or online—are important vehicles for communicating public health messages. Prior studies indicated that the design of these materials significantly affects readers' understanding and retention of these messages [2,3]. In practice, however, too little attention is paid to information selection, wording, formatting, and the use of different modalities (text vs images). The effectiveness of different materials may also depend on the native language and cultural background of the target audience [4,5]. This factor should not be underestimated: according to the American Community Survey [6], 20% of the US population over 5 years of age speak a language other than English at home, and 43.8% of these have limited English proficiency, defined as having a primary language other than English and a limited ability to read, speak, write, or understand English.

Two reasons why health promotion materials are often developed in a cursory manner are time and costs. Typical health promotion research is conducted through surveys, interviews, and focus groups [7,8]. Such studies are costly, time consuming, often subject to selection bias, and of limited generalizability [9], since they most frequently rely on convenience samples of target populations.

Over the last 10 years, the use of Internet surveys has risen in popularity as a result of the ease of access and low costs. Although Internet surveys have raised concerns about generalizability, published studies indicate that Internet surveys that depend on self-selected populations reach more diverse populations than do traditional survey methods that rely on convenience samples. An investigation of a large sample of Internet participants ($n = 361,703$) revealed greater participant diversity in terms of gender, race, age, geographic diversity, and economic status than in traditional studies [10]. Crowdsourcing platforms provide a potential channel for easy access to and recruitment of participants for conducting Internet surveys, especially when trying to access a more diverse population such as non-English speakers. In this paper we

communicate our experience using crowdsourcing technology to test public health promotion materials.

Crowdsourcing

Crowdsourcing is a term used to describe the outsourcing of online tasks performed by a network of people responding to an open call [11]. Although crowdsourcing was initially used for assigning computer-coding tasks, its application has been expanded to product development, advertising, and marketing research [12]. Crowdsourcing, as a method to reach participants, is beginning to revolutionize fields that rely on human experts to perform complex tasks. Collecting data from participant experts has traditionally been difficult, slow, and costly. Examples include speech transcription [13], translations [14], and image labeling [15]. Of the existing online crowdsourcing platforms, the most well known and well studied is Amazon's Mechanical Turk (www.Mturk.com). Mechanical Turk is a crowdsourcing website for brokering so-called human intelligence tasks (HITs)—that is, tasks that are easy for a human to perform but are difficult for a computer. Mechanical Turk connects requesters of HITs to workers and allows for easy task creation, recruitment, compensation, and data collection. The site provides 24/7 access to participants from over 100 countries.

Crowdsourcing in Social Science Research

Since crowdsourcing provides easy, low-cost access to a potentially large pool of participants, it is starting to be considered as a method for research study recruitment in the psychological and behavioral sciences. Mechanical Turk and similar platforms can be used to conduct surveys, opinion polls, or online experiments. Crowdsourcing differs from traditional Internet surveys in that involves limited financial compensation and a pool of regular participants; thus, the recruitment potential, in terms of speed and number of participants, is greatly increased. The potential concern with this method is that there is no face-to-face interaction with participants—the natural question that arises is whether experimental results obtained in this way are valid. Participants might try to “game” the system and not be sufficiently engaged in the task, especially if their primary motivation is monetary compensation. In view of these concerns (see Schmidt [16] for a discussion), several recent studies have analyzed the validity of Mechanical Turk-based data collection for human participants research in political science, psychology, economics, and linguistics [17–21]. These studies have unanimously found that data gathered through Mechanical Turk closely mirrored results collected using standard experimental paradigms, demonstrating a high level of participant engagement, while being significantly easier, faster, and cheaper to obtain.

The relatively low level of compensation typical of Mechanical Turk HITs affected data collection time but not data quality. When compared with convenience samples and other Internet surveys, Mechanical Turk provided data that appeared to be as reliable as other traditional methods [17]. One reason for this outcome may be that, at least for workers located within the United States, the main motivations for working on Mechanical Turk are spending free time in a useful way, having fun, and earning additional income [22,23]. This is different for workers in other countries (eg, India, which hosts the second-largest

segment of Mechanical Turk workers) where compensation is a stronger motivation, though few people use Mechanical Turk as their primary source of income [17]. Berinsky et al [18] and Buhrmester et al [19] also found that Mechanical Turk participants are more diverse and demographically representative than convenience samples.

Crowdsourcing in Public Health

In the context of health care, various forms of crowdsourcing have been used for disaster response [24] and reporting disease outbreaks [25]. These forms of crowdsourcing relied on unpaid volunteers to provide services such as language translation or geolocation. To the best of our knowledge, Mechanical Turk or comparable platforms that implement online microtask brokering have not been studied as a tool in public health communications research or practice, in particular for testing and validating user-oriented health information materials.

Our hypothesis was that crowdsourcing could be an alternative to in-person methods to test public health promotion materials, especially to gain access to non-English-speaking populations. We set out to determine the potential of Mechanical Turk as a rapid, low-cost method for testing the format of health promotion messages designed for diverse populations, including limited English proficiency populations. In particular, we sought to identify the ease of recruitment, costs, and participant demographics associated with using Mechanical Turk to gather rapid and relevant feedback regarding formatting preference from English-speaking and Spanish-speaking individuals for public health communications research.

For this tutorial we chose to focus on our experience studying pediatric dental health messages. Despite being highly preventable, dental disease remains the most common disease of children and adolescents [26]. Tooth decay is four times more common than asthma among adolescents aged 14 to 17 years. Over the past 50 years, major improvements in dental health have been reported nationally in the United States, yet striking disparities remain based on income, age, and race or ethnicity [26]. Many members of society are not informed about, or do not act upon, available dental health messages; therefore, we viewed dental health messaging as an important area for investigating targeted communications.

Methods

Mechanical Turk Setup

Amazon Mechanical Turk facilitates several steps in a crowdsourcing-based study, in particular publishing the task, recruiting participants, collecting the data, and compensating workers. Data preparation and response quality assessment need to be done offline by the researcher.

The first step in setting up a crowdsourcing task is to create and fund a Mechanical Turk account. There is no cost associated with setting up an account, but funds to compensate the workers and to pay the nominal fees charged by the website need to be paid into the account in advance.

The next step involves setting up the task to be performed by workers. The study designer needs to define the overall task,

break it up into microtasks (small tasks that can quickly be performed by an individual worker), formulate the instructions to workers, and prepare the data associated with each task (such as text or images to be annotated, or survey questions). This is done offline, using in-house tools. The Mechanical Turk infrastructure is then used to determine the design of the HIT webpage presented to workers, as well as task and worker attributes, and to upload the data to Mechanical Turk. The desired number and type of HITs are then created automatically from the uploaded data and design template, and they are offered to workers meeting the specified attributes. For example, a researcher might want to annotate 100 different paragraphs. In this case, the template is the form designed to display the paragraph and capture the worker's annotation. Then the data—the 100 different paragraphs—are loaded to create 100 individual HITs.

Task design and attribute specification can be performed using one of three alternatives: the Web-based requester interface, command-line tools, or an application programming interface. In each case, several predefined options are available for the page design (including, for example, checkboxes, drop-down menus, radio buttons, and free-text answers). The task attributes include the compensation per task, number of days the task will be available on Mechanical Turk, the maximum time allotted to any individual worker for completing the task once he or she has accepted it, the number of assignments per task (how many different workers process a given task), and the autoapproval period (the time period after which the results submitted by the worker will automatically be approved). The worker attributes include his or her approval rating (based on previous HITs completed on Mechanical Turk), geographic location, adult content qualification, and any additional qualifications set up by the requester (such as performance on previous tasks by the same requester). The set of worker attributes allows requesters to cultivate pools of trusted workers who habitually deliver good-quality results.

As soon as the template data collection form is created and the data are loaded, researchers can publish the HITs and start receiving answers from workers. Responses can be downloaded, assessed for quality, and approved or rejected online or by uploading a corresponding data file. Once a HIT has been approved, the worker is paid the promised compensation; requesters also have the option of assigning bonuses to workers for particularly satisfying results.

Validity of Responses

One of the main difficulties faced when conducting crowdsourcing studies is assuring the validity of the responses obtained [27]. Since participation is anonymous and linked to monetary incentives, crowdsourcing can attract participants who do not fully engage in the requested tasks or might be unqualified to accurately complete them. There are several ways that a researcher might address this validity issue. The first one is one we already mentioned: setting up qualifications, including qualification tests that need to be passed before the worker can accept a HIT. Second, when the task is associated with an objective ground-truth answer for a subset of the data (such as finding a particular image among a set of images), responses

can be rejected automatically when they do not correspond to the ground truth, and the worker can be blocked. However, this is not possible when the worker's task is to provide a purely subjective assessment. Third, crowd-sourced data collection can involve multiple sequential stages—at each stage, a different set of workers correct the output from previous workers. Fourth, different measures of reliability can be computed on the responses offline, such as outlier statistics or agreement between multiple workers performing the same tasks. Finally, sanity checks (eg, comprehension questions) can be included in the HIT itself.

Crowdsourcing in the Development of Health Promotion Materials

We developed an online survey to test formatting and modality preferences for a variety of messages on pediatric dental health issues (see [Multimedia Appendix 1](#)).

The survey consisted of three sections. In the first part we asked a set of questions about the participants' demographic background, including country of origin, native language, age range, gender, highest education level achieved, whether participants had a regular dentist, and when they last saw a dentist. In the second part, described in more detail below, a paragraph extracted from a pediatric dental education document was presented in four different formats along with text comprehension questions. In the third part participants were asked to select which of the four formats they preferred, followed by an open-text question asking them to state the reasons for their preferences. Optionally, participants were able to provide feedback on the task itself.

In total, we created 12 different survey forms for 12 different documents, each about a different dental health topic. Consent to participate, including information about time to complete the survey and information being collected, was provided prior to initiating the survey. We did not collect any personally identifiable information during the survey; workers are anonymous and only associated with an alphanumeric identification tag. The University of Washington Human Subjects Division approved the study.

For parts 2 and 3 we selected paragraphs from consumer education materials available on US national dental association websites, including the National Institute of Dental and Craniofacial Research, the American Dental Association, the American Academy of Pediatric Dentistry, and the American Academy of Family Physicians. We selected paragraphs to represent a variety of topics regarding childhood dental health, such as tooth brushing, pediatric dental visits, or fluoride use. The content of the selected paragraph was formatted into four versions. Format A consisted only of the running-text paragraph. Format B was a text-only bulleted list. Format C showed the

running-text paragraph and a content-related image (either a photorealistic image or graphics). Format D showed the bulleted list plus the image. All four formats were displayed on the same page. However, the order in which the four formats were presented was determined by random selection. To ensure that participants read and reviewed each of the four versions thoroughly, thus ensuring the validity of their responses, they were requested to answer a different text comprehension question after the presentation of each format. If they answered questions incorrectly, their responses were discarded. We created and tested two versions of the survey, one in English and one in Spanish.

For each survey form, we created a separate HIT on Mechanical Turk. For each HIT, we collected 20 responses (ie, up to 20 different workers answered a single HIT, but a single HIT could not be completed multiple times by the same individual). For each of the two surveys we thus obtained 240 responses.

Participation was limited to individuals located in the United States and those 18 years or older. For the Spanish survey, participants were required to be native Spanish speakers and were asked to specify their country of origin. A separate language qualification test was not applied; however, all Spanish survey materials, including the HIT description and the comprehension questions, were in Spanish, and we did not see any evidence of nonnative speakers taking the Spanish survey. In addition, all comprehension questions were answered correctly. To ensure reliable participants, we also required that they have an approval rate of at least 95% in the HITs they had previously worked on. We allocated 15 minutes for the completion of a single HIT, although we estimated that it could be completed in a much shorter time; the compensation was US \$0.25 per HIT.

Results

The data gathered allowed us to gain insights into participant demographics, the time and costs related to conducting a Mechanical Turk survey, and users' preferences for different messaging formats.

Participant Demographics

We received responses from 236 individual participants for the English survey and 163 for the Spanish survey. Although participation was limited to individuals located in the United States, native Spanish-speaking participants were from 18 different countries. This is a wider distribution over countries than the overall Latino population in the United States, which tends to come more predominantly from Mexico, Central America, and the Caribbean [28]. The five most frequently mentioned countries included three countries in South America ([Table 1](#)).

Table 1. Most frequent countries of origin for native Spanish-speaking participants versus the US Latino population.

Study participants (n = 163)		US Latino population	
Country of origin	%	Country of origin	%
Mexico	20.9%	Mexico	63%
Colombia	9.4%	Puerto Rico	9.2%
Argentina	9.2%	Cuba	3.5%
Peru	6.1%	El Salvador	3.3%
El Salvador	4.6%	Dominican Republic	2.8%

Table 2 summarizes key demographic characteristics and the time required to answer our HITs. The overall demographic composition of our respondent populations is similar to compositions observed in previous studies [22,29]—that is,

Mechanical Turk workers were predominantly young and well educated. It is noteworthy that the Spanish-speaking respondents overall seem to have been even more highly educated than their English-speaking counterparts.

Table 2. Comparison of Spanish-speaking and English-speaking survey participants.

Demographic characteristic	Spanish-speaking (n = 163)	English-speaking (n = 236)	P value
Time to completion of HITs ^a	~12 days	~6 days	NA ^b
Females, n (%)	73 (45%)	138 (58.5%)	.007
Age 18–40 years, n (%)	130 (79.8%)	184 (78%)	.67
College or graduate degree, n (%)	117 (71.8%)	102 (43.2%)	<.001
Average time/HIT (minutes:seconds), mean (SD)	5:20 (2:50)	4:05 (3:31)	<.001

^a Human intelligence tasks.

^b Not applicable.

Time and Cost

The responses for the English survey were collected within 6 days; the Spanish survey took approximately twice as long. The total cost including the Mechanical Turk commission amounted to US \$374.

Preliminary Preferences and Feedback

The main goal of this study was to explore the feasibility of using crowdsourcing to obtain feedback on information

presentation options. Thus, we mention the actual results regarding respondents' preferences primarily for the sake of completeness. Both the Spanish and English survey results indicate that participants largely preferred the format that included bulleted text with an image related to the text. The remaining preferences were evenly distributed among the other three formats (**Table 3**).

Table 3. Comparison of format preferences Spanish-speaking and English-speaking survey participants (total n = 399).

Format	Spanish-speaking (n = 163)		English-speaking (n = 236)	
	n	%	n	%
Bullet + image	106	65.0%	136	57.6%
Bullet alone	22	14%	46	20%
Paragraph with images	27	17%	33	14%
Paragraph alone	7	4%	21	9%

In addition to asking participants to answer the survey, we gave them the opportunity to comment on the survey, which was extremely useful. Several participants provided relevant feedback about the HITs and the value of the health promotion documents. In general, participants seemed to enjoy the task and mentioned that they found it educational or that they thought

it was a useful research study. As an example, one participant commented:

My kids' adult teeth got ruined from giving them bottles at bedtime and it pooling in their mouth. Wish I would have known this back when my kids were little.

Other participants mentioned that they disliked photorealistic images of dental diseases but that they valued the message (translated from Spanish):

The image is too graphic but it is adequate to convey the message.

Although the picture is disgusting, it is real, and it is important so moms can be aware of the consequences of sugar in children's beverages.

Others commented on specific issues of the Spanish wording, preferring expressions from some dialects of Spanish (eg, Mexican) over others.

Discussion

Crowdsourcing is a new method for gathering data when human participation is required. Our results show that this technology can be used for gathering useful feedback on the design of health information materials from a large number of participants in a rapid and inexpensive manner, which is in line with results from previous studies in related fields [17-21]. This is particularly important for public health purposes, where materials might not be tested because of time and resource constraints, making it difficult to access a sufficiently large subset of the intended target audience. For example, public health agencies in the United States are bound by federal requirements to provide certain health information to non-English-speaking populations [30], which makes having access to such individuals an important step in adapting health promotion materials. Through crowdsourcing we were able to rapidly recruit a large number of native Spanish speakers, residing in the United States, originating from a large number of countries. In addition to collecting data about preferred formats, we also received feedback regarding the survey itself and useful suggestions from Spanish speakers regarding alternative vocabulary and culturally appropriate images.

An unanticipated side effect of this study was the communication of dental health promotion messages to participants who stated they had previously been unaware of some of the dental health recommendations. Several participants noted their appreciation for participating in an educational HIT. It is conceivable that crowdsourcing can be used not only to test and develop health messages but also to distribute messages to targeted populations. In the ideal case, both purposes could be combined in one crowdsourcing application.

The preliminary results of our survey suggested that both English-speaking and Spanish-speaking participants in the United States preferred a format with bullets and images. The preference for images is consistent with results from prior studies using more traditional methods [4,31,32]. Despite this overall trend, preferences were not unanimous: 17.7% of Spanish speakers and 28.4% of English speakers preferred formats without an image. This variation suggests that in the future, feedback regarding individual messaging formats preferences could be used to inform message tailoring. In addition, some Spanish-speaking participants provided feedback on the wording of the survey, which suggests that crowdsourcing

could be used to obtain user feedback to edit terms or phrasing in future versions of the survey.

Limitations of Crowdsourcing

In line with previous studies, the demographic information collected in this study indicates that the population we reached was younger and more educated than the general population. Reaching primarily younger participants could be helpful when targeting messages pertinent to young parents or young adults in general (eg, on sexually transmitted diseases, drug and alcohol abuse, and injury prevention). Interestingly, Spanish speakers overall reported higher levels of education than English speakers. Clearly, one limitation of the Mechanical Turk recruitment method is the difficulty in reaching populations with low literacy, low computer skills, low educational level, or the elderly, although those are often the populations most in need of health information and support. However, 6%–7% of both the Spanish-speaking and English-speaking participants together had less than high school education. It is possible that, with a sufficiently large sample, valid information about some of the less well-represented demographic groups could be obtained at costs that are still lower than those of traditional surveying methods. Alternative methods to Internet surveying may include crowdsourcing through cell phone text messaging, which may later prove to be an effective way to gain feedback from other hard-to-reach groups.

Given the monetary compensation of Mechanical Turk, it is possible that individuals might try to game the system by answering questions quickly without serious consideration. To guard against this, we required participants to accurately answer questions about the content before the HIT would be accepted. Although there is no guarantee that participants gave true answers, all participants answered the content questions accurately, and their comments suggested that the participants considered their answers carefully.

Limitations of this Study

The main limitation of this study was that, because it was a pilot study, we did not solicit a statistically significant number of participants to draw conclusions about the responses regarding formatting preferences. Our goal in this study was to investigate how easy or difficult it would be to gain access to specific populations through crowdsourcing. An expanded study is needed to identify whether there are significant differences in messaging preferences between English-speaking and Spanish-speaking populations. Our results suggest that by using Mechanical Turk we will be able to recruit a large sample of participants in a relatively short time and at low costs. In addition, we did not compare the results with more conventional survey methods using convenience samples. However, Mechanical Turk has been compared with traditional experimental paradigms in several previous studies (see above) and has been validated as a way to gather survey responses; moreover, we envision crowdsourcing as a different way to access participants that should not necessarily be compared with conventional survey methods. Our results suggest that the demographics of the surveyed population using crowdsourcing are likely to be *different* from those of convenience samples accessed through more traditional methods. As a result,

Conclusions

We used crowdsourcing to recruit a substantial number of English-speaking and Spanish-speaking participants for a survey on health promotion materials in 2 to 4 days for low costs. Results suggest that crowdsourcing could become a valuable research tool in public health communications research.

This research was funded by the University of Washington School of Public Health Royalty Research Fund Grant, the National Library of Medicine Grant 1R01LM010811-01, and Dr Capurro's Fulbright-MECESUP Faculty Development Grant. The authors wish to thank Kate Cole for review and preparation of the manuscript.

None declared.

Sample task performed by study participant.

[PDF File (Adobe PDF File), 1MB - jmir_v14i3e79_app1.pdf]

1. McHorney CA, Kosinski M, Ware JE. Comparisons of the costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: results from a national survey. Med Care 1994 Jun;32(6):551-567. [Medline: [8189774](#)]
2. Maibach E, Parrott R. Designing Health Messages: Approaches From Communication Theory and Public Health Practice. Thousand Oaks, CA: Sage Publications; 1995.
3. Knapp P, Raynor DK, Jebar AH, Price SJ. Interpretation of medication pictograms by adults in the UK. Ann Pharmacother 2005;39(7-8):1227-1233. [doi: [10.1345/aph.1E483](#)] [Medline: [15897266](#)]
4. Houts PS, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006 May;61(2):173-190. [doi: [10.1016/j.pec.2005.05.004](#)] [Medline: [16122896](#)]
5. Ayala GX, Elder JP, Campbell NR, Engelberg M, Olson S, Moreno C, et al. Nutrition communication for a Latino community: formative research foundations. Fam Community Health 2001 Oct;24(3):72-87. [Medline: [11563946](#)]
6. American Community Survey. US Census Bureau. 2009. Language Spoken at Home URL: http://factfinder.census.gov/servlet/SSTable?_lang=en&_ss=ACSR&_ds=ACS2009NRC&_tid=ACSR&_cid=ACSR&_all_geo_types=N&_geo_id=0001&_sex=&_race=&_hispanic=&_age=&_mar=&_nat=&_anncode=0001 [accessed 2011-09-19] [WebCite Cache ID 61owyreJ4]
7. Bartholomew LK. Planning Health Promotion Programs: An Intervention Mapping Approach. San Francisco, CA: John Wiley and Sons; 2006.
8. Hubley J, Copeman J. Practical health promotion. Cambridge: Polity; 2008.
9. Neuman WL. Social Research Methods: Qualitative and Quantitative Approaches. 6th edition. Boston, MA: Pearson/Allyn and Bacon; 2006.
10. 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](#)]
11. Howe J. The rise of crowdsourcing. Wired 2006;14(6).
12. Whitla P. Crowdsourcing and its application in marketing activities. Contemp Manag Res 2009;5(1):15-28.
13. Marge M, Banerjee S, Rudnicky AI. Using the Amazon Mechanical Turk for transcriptions of spoken language. In: Proceedings. 2010 Presented at: International Conference on Acoustics, Speech and Signal Processing; Mar 14-19, 2010; Dallas, TX, USA.
14. Zaidan OF, Callison-Burch C. Crowdsourcing translation: professional quality from non-professionals. In: Proceedings. 2011 Presented at: 49th Annual Meeting of the Association for Computational Linguistics; June 19-24, 2011; Portland, OR, USA.
15. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: Proceedings. 2009 Presented at: IEEE Computer Science Conference on Computer Vision and Pattern Recognition; June 20-25, 2009; Miami Beach, FL, USA.

16. Schmidt L. Crowdsourcing for human subjects research. 2010 Presented at: CrowdConf; Oct 4, 2010; San Francisco, CA, USA.
17. Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk. *Behav Res Methods* 2012 Mar;44(1):1-23. [doi: [10.3758/s13428-011-0124-6](https://doi.org/10.3758/s13428-011-0124-6)] [Medline: [21717266](https://pubmed.ncbi.nlm.nih.gov/21717266/)]
18. Berinsky AJ, Huber GA, Lenz GS. Using Mechanical Turk as a subject recruitment tool for experimental research. *Political Analysis* 2012 (forthcoming).
19. Buhrmester M, Kwang T, and Gosling SD. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspect Psychol Sci* 2011;6(1):3-5.
20. Horton JJ, Rand DG, Zeckhauser RJ. The online laboratory: conducting experiments in a real labor market. *Exp Econ* 2011;14(3):399-425. [doi: [10.1007/s10683-011-9273-9](https://doi.org/10.1007/s10683-011-9273-9)]
21. Sprouse J. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behav Res Methods* 2011 Mar;43(1):155-167. [doi: [10.3758/s13428-010-0039-7](https://doi.org/10.3758/s13428-010-0039-7)] [Medline: [21287108](https://pubmed.ncbi.nlm.nih.gov/21287108/)]
22. Ross J, Irani L, Silberman MS, Zaldivar A, Tomlinson B. Who are the crowdworkers?: Shifting demographics in Mechanical Turk. In: *Proceedings. 2010 Presented at: 28th ACM Conference on Human Factors in Computing Systems*; Apr 10-15, 2010; Atlanta, GA, USA.
23. Kittur A, Chi EH, Suh B. Crowdsourcing user studies with Mechanical Turk. 2008 Presented at: 26th Annual SIGCHI Conference on Human Factors in Computing Systems; Apr 5-10, 2008; Florence, Italy.
24. Gao H, Barbier G, Goolsby R. Harnessing the crowdsourcing power of social media for disaster response. *IEEE Intell Syst* 2011;26(3):10-14. [doi: [10.1109/MIS.2011.52](https://doi.org/10.1109/MIS.2011.52)]
25. Freifeld CC, Chunara R, Mekaru SR, Chan EH, Kass-Hout T, Ayala Iacucci A, et al. Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS Med* 2010;7(12):e1000376 [FREE Full text] [doi: [10.1371/journal.pmed.1000376](https://doi.org/10.1371/journal.pmed.1000376)] [Medline: [21151888](https://pubmed.ncbi.nlm.nih.gov/21151888/)]
26. Satcher D. US Department of Health and Human Services, National Institute of Dental and Craniofacial Research, National Institutes of Health. 2000. Oral Health in America: A Report of the Surgeon General (Executive Summary) URL: <http://www.nidcr.nih.gov/DataStatistics/SurgeonGeneral/Report/ExecutiveSummary.htm> [accessed 2012-05-15] [WebCite Cache ID 67gKQhBaY]
27. Downs JS, Holbrook MB, Sheng S, Cranor LF. Are your participants gaming the system? Screening Mechanical Turk workers. 2010 Presented at: 28th International Conference on Human Factors in Computing Systems (CHI '10); Apr 19-15, 2010; Atlanta, GA, USA.
28. Lopez MH. U.S. Hispanic country of origin counts for nation, top 30 metropolitan areas. Washington, DC: Pew Hispanic Center; 2011 May 26. URL: <http://pewhispanic.org/reports/report.php?ReportID=142> [accessed 2011-09-19] [WebCite Cache ID 61p3ma71v]
29. Ipeirotis P. Social Science Research Network. 2010 Mar. Demographics of Mechanical Turk URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1585030 [accessed 2012-03-23] [WebCite Cache ID 6605eIcR9]
30. Department of Health and Human Services. Guidance to federal financial assistance recipients regarding Title VI prohibition against national origin discrimination affecting limited English proficient persons. *Federal Register* 2003;68(153):47311-47323.
31. US Department of Health and Human Services, Office of Disease Prevention and Health Promotion. 2011. Quick Guide to Health Literacy: Strategies: Improve the Usability of Health Information URL: <http://www.health.gov/communication/literacy/quickguide/Quickguide.pdf> [accessed 2011-09-19] [WebCite Cache ID 61p4IpHge]
32. Michielutte R, Bahnson J, Dignan MB, Schroeder EM. The use of illustrations and narrative text style to improve readability of a health education brochure. *J Cancer Educ* 1992;7(3):251-260. [doi: [10.1080/08858199209528176](https://doi.org/10.1080/08858199209528176)] [Medline: [1419592](https://pubmed.ncbi.nlm.nih.gov/1419592/)]

Abbreviations

HIT: human intelligence task

Edited by G Eysenbach; submitted 23.01.12; peer-reviewed by J DeShazo, H Spallek; comments to author 09.03.12; revised version received 23.03.12; accepted 26.03.12; published 04.06.12.

Please cite as:

Turner AM, Kirchhoff K, Capurro D

Using Crowdsourcing Technology for Testing Multilingual Public Health Promotion Materials

J Med Internet Res 2012;14(3):e79

URL: <http://www.jmir.org/2012/3/e79/>

doi: [10.2196/jmir.2063](https://doi.org/10.2196/jmir.2063)

PMID: [22664384](https://pubmed.ncbi.nlm.nih.gov/22664384/)

©Anne M. Turner, Katrin Kirchhoff, Daniel Capurro. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 04.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Review

Health and Wellness Technology Use by Historically Underserved Health Consumers: Systematic Review

Enid Montague¹, PhD; Jennifer Perchonok¹, MS

Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:

Enid Montague, PhD

Department of Industrial and Systems Engineering

University of Wisconsin-Madison

3017 University Avenue

Madison, WI, 53711

United States

Phone: 1 608 890 3546

Fax: 1 608 262 8454

Email: emontague@wisc.edu

Abstract

Background: The implementation of health technology is a national priority in the United States and widely discussed in the literature. However, literature about the use of this technology by historically underserved populations is limited. Information on culturally informed health and wellness technology and the use of these technologies to reduce health disparities facing historically underserved populations in the United States is sparse in the literature.

Objective: To examine ways in which technology is being used by historically underserved populations to decrease health disparities through facilitating or improving health care access and health and wellness outcomes.

Methods: We conducted a systematic review in four library databases (PubMed, PsycINFO, Web of Science, and Engineering Village) to investigate the use of technology by historically underserved populations. Search strings consisted of three topics (eg, technology, historically underserved populations, and health).

Results: A total of 424 search phrases applied in the four databases returned 16,108 papers. After review, 125 papers met the selection criteria. Within the selected papers, 30 types of technology, 19 historically underserved groups, and 23 health issues were discussed. Further, almost half of the papers (62 papers) examined the use of technology to create effective and culturally informed interventions or educational tools. Finally, 12 evaluation techniques were used to assess the technology.

Conclusions: While the reviewed studies show how technology can be used to positively affect the health of historically underserved populations, the technology must be tailored toward the intended population, as personally relevant and contextually situated health technology is more likely than broader technology to create behavior changes. Social media, cell phones, and videotapes are types of technology that should be used more often in the future. Further, culturally informed health information technology should be used more for chronic diseases and disease management, as it is an innovative way to provide holistic care and reminders to otherwise underserved populations. Additionally, design processes should be stated regularly so that best practices can be created. Finally, the evaluation process should be standardized to create a benchmark for culturally informed health information technology.

(*J Med Internet Res* 2012;14(3):e78) doi:[10.2196/jmir.2095](https://doi.org/10.2196/jmir.2095)

KEYWORDS

health care disparities; biomedical technology; health education; health knowledge, attitudes, and practice; health care quality, access, and evaluation; educational technology; cultural diversity

Introduction

While the visibility of health disparities has recently come to the forefront of the US health care agenda, the topic of health

care disparities is not new. In 1984 the health of the nation was addressed in the “Health, United States, 1983” report conducted by the US Department of Health and Human Services. This report stated that African Americans and other racial and ethnic minorities were experiencing a higher burden of death and

illness than the rest of the nation [1]. As a reaction to this report, the Secretary of the Department of Health and Human Services created the first group solely designated to study minority health issues—the Task Force on Black and Minority Health. In 1985, this group published a comprehensive study on minority health problems, “Report of the Secretary’s Task Force on Black and Minority Health.” The report brought more awareness of health disparities in historically underserved populations and spurred research [1].

In the United States, historically underserved populations are growing in size, and hence health disparities are affecting a growing proportion of Americans. For instance, while 2000 census findings showed that 82% of the population was white, by 2015 this number is predicted to decrease to 79%. At that time, it is expected that these will be 5% Asian, 13% African American, and 15% Latino [2]. By 2050, ethnic populations will double in size in the United States and constitute 40% of the population [2]. Similar studies conducted outside of the census found similar results. Partida reported that one in eight Americans is foreign born, and 45% of children less than 5 years of age are not white [3]. Beyond ethnicity, the percentage of older Americans is also increasing, with the oldest (85+ years of age) and ethnic elderly populations growing at the fastest rates [4].

With historically underserved populations growing in the United States, it is important to study the potential and existing health disparities facing them. While there is no consensus regarding the specific definition of what constitutes a health disparity, the National Institutes of Health defined a health disparity as “differences in the incidence, prevalence, mortality, and burden of diseases and other adverse health conditions that exist among specific population groups in the United States” [5]. More specifically health disparities include inequalities and inequities in (1) environment, (2) access to, use of, and quality of care, (3) health status, or (4) a specific health outcome [5]. Examples of racial and ethnic health disparities include certain populations with exceedingly high rates of cardiovascular disease, diabetes, asthma, and cancer [6]. While factors including lower socioeconomic status, lack of insurance, lower levels of education, and living in communities with more environmental hazards have been cited as social determinants outside of the health care system, these do not fully account for health disparities [6]. A patient’s culture may contribute to the disparities facing them by influencing health beliefs, values, preferences, and behaviors. For instance, a patient’s ability to recognize symptoms and then effectively describe the symptoms to a provider will influence their interactions with their provider, which can in turn affect their health outcomes [6]. The United States made eliminating health disparities one of its main goals through Healthy People 2010, a federal interagency workgroup that provides 10-year national objectives for improving the general health of Americans [7]. Healthy People 2010 targeted disparities based on race and ethnicity, gender, education, income, geographic location, disability status, or sexual orientation [8].

As technology is being used to further the success of the health care system, there is interest in understanding how the increased use of technology affects the already unequal ability of

minorities to access health care [9]. Health technology has been used since as early as the middle to late 19th century, when electrocardiograph data were transmitted over telephone wires [10]. Today, health information technology (IT) is used to benefit both the health care consumer and public health as a whole. Health care consumers benefit from health IT by receiving a higher quality of care, reduction in medical errors, fewer duplicate treatments and tests, decreases in paperwork, lower health care costs, access to health information, and access to affordable care [11]. The public health sector benefits from health IT, as it can facilitate earlier detection of infectious disease outbreaks, improve the tracking of chronic disease management, and gather de-identified data for research [11].

Technology can be used in a variety of ways to positively affect historically underserved health care consumers. For example, telemedicine has been suggested as a possible way to address health care disparities among historically underserved urban populations. Research shows that urban communities are often unable to access health care in a timely manner due to low physician-to-population ratios, limited specialty care, and overcrowded, inadequate, and inefficient organizational structures [12]. Telemedicine is an innovative way to decrease the health care gap through mitigating geographic barriers [12].

To promote widespread adoption of health IT, the US Department of Health and Human Services established the Office of the National Coordinator for Health Information Technology [13]. Health IT refers to “a variety of electronic methods used to manage information about people’s health and health care, on both an individual and a group level” [14]. Research has shown that health IT can enhance quality, communication [15], and cost-effective care [16], and can facilitate culturally competent outreach and education [17].

The purpose of this review was to examine ways in which technology is being used by historically underserved populations in order to decrease the health disparity through facilitating or improving health care access and health and wellness outcomes. While several studies have investigated how historically underserved populations use technology when addressing their health, these studies focused on a single historically underserved group or a single health issue. We used a methods-description approach method to synthesize published research from reference databases to draw a larger conclusion from the current literature [18]. We explored four main questions from the reviewed papers. (1) Which types of technologies are used to address potential disparities? (2) Which health issues are addressed in the reviewed papers? (3) Which historically underserved groups are targeted for technology-based interventions? (4) How are the health benefits and technologies evaluated? The systematic review was conducted in four reference databases (PubMed, PsycINFO, Web of Science, and Engineering Village) with search strings consisting of three topics: technology, historically underserved groups, and health. Findings are divided into five sections, each answering one of the five main questions. Outcomes include recommendations for increased use of certain technology, along with recommendations to use culturally informed technology in regard to distinct types of health conditions.

Methods

Definitions

The term minority has been used often in health research. However, the term is problematic, as it can create a sense of inferiority for the population in question [19]. Eysenbach suggested that eHealth is a broad topic encompassing 10 main concepts, one of which is equity [20]. He noted that certain patient populations are disadvantaged based on their lack of money, skills, and access to computers. However, the use of the term minority only further perpetrates these inequities. For this reason, we will not use the term minority in this paper. Instead, we will use the term historically underserved to refer to populations that are disadvantaged based on their race, ethnicity, age, gender, socioeconomic status, health status, or geographic location.

Larson stated that simplistic definitions of health should be avoided, as they lead to simplistic measures of health, health outcomes, and quality of care [21]. Therefore, it is important to use a more holistic definition of health that includes wellness; for instance, the World Health Organization defined health as “a state of complete physical, mental and social well-being, not just absence of disease” [21]. Pervasive health care takes this concept a step further and can be defined as “healthcare to anyone, anytime, and anywhere by removing locational, time and other restraints while increasing both the coverage and the quality of healthcare” [22].

The *Health Technology Assessment Handbook* defined health technology as “a collective term for procedures and methods for examination, treatment, care and rehabilitation of patients, including instruments, drugs, and preventive procedures” [23]. Health IT, which differs slightly from health technology, refers to the implementation of information processing that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge to facilitate both decision making and communication [24]. Health IT used directly by consumers is called consumer health IT. Or and Karsh defined consumer health IT as “computer-based systems that are designed to facilitate information access and exchange, enhance decision making, provide social and emotional support, and help behavior changes that promote health and well-being” [25].

eHealth refers to the use of electronic communication and information technology within the health sector. Tools often referred to in connection with eHealth include personal digital assistants, compact discs and DVDs, and interactive games [26]. Telemedicine, which is a part of eHealth, allows providers and patients in different geographic locations to communicate through computers, information, and telecommunication [12]. Telehealth, which is often used synonymously with telemedicine, is defined by the World Health Organization as telemedicine used by others beyond the physician [10] such as nurses and pharmacists. For this review, the author of the original paper differentiated between these 2 terms. For instance, if the author of the paper under review used the term telemedicine, we used it for this review; if the author of the

reviewed paper used telehealth, we used it for this review as well.

Electronic health records (EHRs) are an electronic form of the traditional patient health record (patient’s health profile, and environmental and behavioral information). EHRs include a time dimension and allow multiple providers to contribute information to the record [27]. EHRs have been shown to have a positive influence on quality of care, patient safety, and system delivery [17]. Electronic medical records are similar to EHRs except they are created solely for care delivery organizations—that is, hospitals and physician’s offices [28]. EHRs have the ability to increase access to health care, reduce medication errors, and improve administrative efficiency and quality of care [16]. In contrast to the EHR, personal health records are an optional tool that allows people to manage their own health records [29]. The personal health record is a lifelong resource of health information that is managed by the individual in an electronic, universally available form [29].

For this review, we used a broad definition of technology that includes technology designed for both health and wellness. In this review, health-specific technology designed specifically for the clinical setting includes health IT, EHRs, and telemedicine. We also included wellness informatics, defined as “a human-centered computing science focused on the design, deployment, and evaluation of human-facing technological solutions to promote and manage wellness acts such as the prevention of disease and the management of health” [30], in this review. Wellness informatics encompasses technology that may have little or no interaction with the health care system and is used primarily by the consumer [30]. For this review, wellness informatics tools included media technology created for other domains, such as television, radio, and computers.

Search Strategy

From July to October 2011, we searched the online reference databases PubMed, Web of Knowledge, PsycINFO, and Engineering Village. For each database, we chose keywords to match the specific database’s thesaurus and used them to create search phrases. Each search phrase consisted of three key components: a word or phrase considering historically underserved populations, a word or phrase considering technology, and a phrase considering health, health access, or wellness (Table 1). Keywords about historically underserved populations included cultural diversity, ethnic groups, minority groups, cultural competency, ethnocentric, cross-cultural difference, racial and ethnic attitudes, racial and ethnic differences, and racial and ethnic discrimination. Keywords pertaining to health or health access included health education, patient acceptance of health care, attitudes to health, access to information, electronic health care, health, health system, and patient care. Finally, words considering technology included telemedicine, technology, medical technology, educational technology, medical information systems, eHealth, and health technology. When combined into the longer 3-part phrases, a total of 424 search phrases were used.

Table 1. Search terms by topic.

Historically underserved populations	Technology	Health and health access
Cultural diversity ^a	Telemedicine ^{a,b,c}	Health education ^a
Ethnic groups ^{a,d}	Technolog ^{a,b,c}	Patient acceptance of healthcare/ethnology ^a
Medically underserved areas ^a	Medical technolog ^{a,d}	Acceptance of healthcare ^a
Minority group ^a	Educational technology ^a	Attitudes to health ^a
Cross-sectional studies ^a	Electronic healthcare ^d	Access to information ^a
Cultural competenc ^{a,d}	E-health ^d	Health knowledge, attitudes, practice ^a
Health status disparities ^a	Health technolog ^d	Evaluation ^d
Disparit ^d	Healthcare technolog ^d	Health access ^d
Social factors ^d	Medical information systems ^b	Technolog* acceptance ^d
Ethnocentric ^d	Medical computing ^b	Healthcare professionals ^d
Reference group culture ^d	Information technology ^{b,c,d}	Health system ^d
Cultur* bias ^d		Health ^d
Minorit ^d		Healthcare ^b
Cultural aspects ^b		Patient care ^b
Culture bound syndromes ^c		Health disparities ^c
Ethnology ^c		Health attitudes ^c
Cross cultural differences ^c		Health knowledge ^c
Racial and ethnic attitudes ^c		Health impairments ^c
Racial and ethnic differences ^c		Health complaints ^c
Racial and ethnic groups ^c		
Race and ethnic discrimination ^c		

^a PubMed.^b Engineering Village.^c PsycINFO.^d Web of Science.

Inclusion and Exclusion Criteria

The scope of the review was focused by establishing inclusion and exclusion criteria. The selection criteria were that the paper (1) focused on a specific priority population(s), (2) discussed how the populations' identity affected their experience within the health care system, and (3) discussed how technology use affected the experience.

We excluded studies if they (1) were published over 15 years ago (prior to 1996), (2) were not in English, (3) were conducted outside of the United States, (4) did not deal with health or wellness, (5) discussed mental health, end-of-life care, or dental care, or examined cost as the main variable, and (6) discussed the historically underserved population as a current or future employee of the health system instead of as a patient.

Analysis

We used a methods-description approach to analyze papers that met the inclusion criteria. This method documented the objective characteristics (as they were described by the primary author) of each study's methods [18]. In compliance with the methods-description approach and to ensure standardized data extraction of the reviewed papers, we created a data table [18] with the following sections: title, author, purpose, and key findings. After completing the table, we defined recurrent topics through coding. Coding is defined as the "analytical process through which concepts are identified and dimensions are discovered in data" [31]. Through use of coding, the following ideas were explored: the targeted historically underserved group, the health issue examined, how technology was used, evaluation techniques, and barriers to access or adoption ([Multimedia Appendix 1](#) [4,9,12,13,15-17,32-149]).

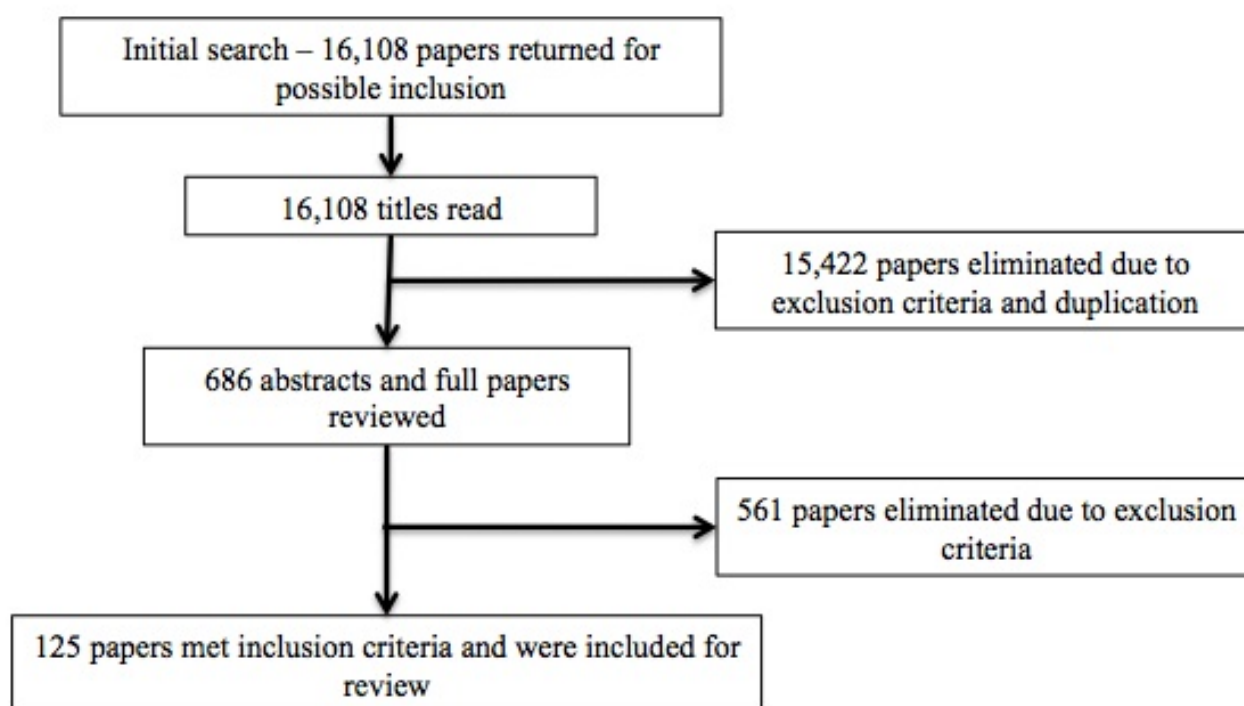
Results

The 424 search phrases returned 16,108 papers. We excluded 15,422 papers as duplications or via the exclusion criteria through reading the titles. After reviewing the abstracts and full papers, we eliminated another 561 papers as not meeting the inclusion criteria. A total of 125 papers met the inclusion criteria and were included in this review (Figure 1). An overview of the 125 papers can be found in [Multimedia Appendix 1](#).

All selected papers discussed the health disparity facing the historically underserved group in question and the importance of closing the gap or reducing the disparity. One-quarter of the papers (32) focused only on the health disparity without analyzing a potential solution. The remaining 93 papers briefly

mentioned the disparity but focused on accessing a possible solution to lessen the disparity. For instance, 1 report discussed the health disparities facing the Hispanic community in the background of the paper. However, the main purpose of the paper was to determine the effectiveness of *La Clínica del Pueblo*, a health education collaboration that uses radio to increase medical knowledge and have a positive affect on health behaviors [44]. The design and development of the technology was discussed in detail in only 13 of the reviewed papers. More often, the authors simply stated that technology was used in an attempt to lessen the disparity. Additionally, 5 papers were review papers; of the 5 review papers, 2 discussed diabetes [36,62], 2 discussed general health and health IT [9,15], and 1 discussed health literacy for people whose second language is English [65].

Figure 1. Flow diagram of the study selection process.



Which Types of Technologies are Used to Address Negative Health Outcomes for Historically Underserved Populations?

We identified 30 types of technology in the selected papers (Table 2). The technology included both health informatics tools (personal digital assistant, radio, Internet, telephone, mobile computer, mobile phone, videotape, computer, kiosk, MP3, television, compact disc, multimedia tool, instant messaging, and fax machine) and more traditional health technology

(general health IT, medical technology, telemedicine, telehealth, telemanagement, electronic medical records, personal health records, EHR, eHealth, assisted reproductive technology, high technology hospitals, rapid human immunodeficiency virus (HIV) testing, implantable cardioverter defibrillator, cochlear implants, and assistive technology). The technology was used in a variety of ways, including as educational tools, as pieces of interventions, or as collaboration tools between physicians and patients.

Table 2. Paper breakdown by technology.

Technology	Number of papers
Video	34
Internet (email, social networking sites)	23
Telemedicine	10
Computer (computers in clinics)	9
Television (advertisements and shows)	8
General health information technology	6
Electronic health record	6
Radio	5
Telephone	5
Mobile phone (text messaging)	5
Assisted reproductive technology	5
Multimedia tool	4
Assistive technology	4
Telehealth	3
Compact disc	2
Kiosk	2
Telemanagement	2
eHealth	2
Medical technology	1
Electronic medical record	1
Personal health record	1
Personal digital assistant	1
Mobile computer	1
High-technology hospitals	1
MP3	1
Rapid HIV ^a testing	1
Implantable cardioverter defibrillator	1
Cochlear implants	1
Instant messaging (on a computer)	1
Fax machine	1

^a Human immunodeficiency virus.

At Which Disease, Health Problem, or Potential Problem is the Technology Aimed?

The reviewed papers discussed 23 health issues (Table 3). Roughly one-quarter of the papers (33) did not focus on a single health topic, but instead discussed the general health of a population. Other papers examined health issues such as disease management (eg, diabetes, asthma, and obesity); health behaviors (eg, nutrition and smoking); and short-term issues (eg, breast-feeding, issues facing pregnant mothers, and child development).

Nearly half of the papers (62) examined the use of technology to create effective and culturally informed interventions (16

papers) or educational tools (46 papers). The reviewed papers pointed to many interventions and educational tools that were successfully designed for a historically underserved group. A study found that having famous athletes, musicians, and other celebrities from the African American community record commercials for adolescents' MP3 players resulted in better health knowledge about asthma [95]. Another study created a telenovela for Latinas to discuss breast cancer in the dramatic and narrative format of a typical telenovela. Relating to the women on a cultural level, such as through the telenovela, resulted in higher breast cancer knowledge for the participants [144].

Table 3. Paper breakdown by health issue.

Health issue	Number of papers
General health	33
Cancer	17
Diabetes	14
HIV/AIDS ^a	14
Nutrition, physical activity	8
Sexually transmitted infections	7
Reproduction	5
Obesity	4
Cardiovascular disease, heart problems	4
Breast-feeding	3
Smoking	3
Asthma	3
Persons with disabilities	3
Pregnancy issues	2
Pharmacy	2
Sensorineural hearing loss	1
Organ donation	1
Hepatitis C	1
Health literacy	1
High blood pressure	1
Poison control	1
Hypertension	1
Child development	1

^a Human immunodeficiency virus/acquired immunodeficiency syndrome.

Which Historically Underserved Groups are Technology-Based Interventions Designed for in the Literature?

The papers identified 19 different historically underserved populations (Table 4). Of the reviewed papers, 18 discussed multiple groups (eg, elderly African Americans or Hispanic women) and therefore appear in 2 categories in the table. In 8

the group in question self-identified as “racial and ethnic minorities.” We copied this term in this review paper only when the original author did not provide sufficient details to determine which racial or ethnic minorities were being examined. In addition to racial and ethnic minorities, the reviewed papers also included historically underserved groups that were characterized by their age, gender, location, and socioeconomic status.

Table 4. Paper breakdown by historically underserved group.

Historically underserved group	Number of papers
African American	64
Hispanic	51
Women (mothers)	26
Low socioeconomic status	11
Elderly	11
Adolescents, teens, and children	8
Racial and ethnic minorities	8
English as a second language	5
Native American and Alaskan	4
Men	4
Rural	4
Underresourced setting, underserved community	3
Community health center: underserved, low socioeconomic status, racial and ethnic groups	2
People getting tested for HIV ^a	2
Asian American	1
Immigrant	1
Homeless	1
People with AIDS ^b	1
People living with HIV	1

^a Human immunodeficiency virus.

^b Acquired immunodeficiency syndrome.

How Are the Health Benefits and Technologies Evaluated?

Other than the 3 review papers, the papers all used formative technology evaluation. They used two forms of evaluation: (1) evaluation of health changes related to use of the technology and (2) evaluation of the technology itself; few papers (23) used both types of evaluation. A total of 76 of the papers evaluated health changes due to use of the technology (eg, changes in health knowledge, health behavior changes, biometric changes, or changes in health-related quality of life). Of the 107 papers that evaluated technology, 57 evaluated acceptance of the technology (satisfaction or acceptance, usefulness, and willingness to use), 14 evaluated usability (ease of use), 35 evaluated the user's ability to access the technology (access or usage rates and number of websites or television advertisements

with the desired information), and 1 measured improvements in technology literacy. In addition, 64 papers relied on the participants' self-report to evaluate the technology, 14 measured ease of use, 22 measured usefulness of the technology, and 28 evaluated satisfaction with the technology. When an intervention or educational tool was evaluated, some of the authors (25 papers) measured improvement in participant health knowledge, while others measured behavior change (22 self-reported behavior changes and 18 observed behavior changes). Furthermore, 10 papers measured biometric changes in the observed health condition, 31 examined access and usage rates of the technology, and 7 recorded whether patients were interested in using the technology in the future. Finally, 4 papers measured the number of websites or television advertisements viewed by the population being studied. [Table 5](#) lists the evaluation methods.

Table 5. Evaluation metrics.

Evaluation metric	Number of papers
Evaluation of health changes related to use of the technology	
Health knowledge	25
Behavior change (self-reported)	22
Behavior change (observed)	18
Biometric change	10
Health-related quality of life	1
Evaluation of the technology itself	
Access and usage rates	31
Self-reported satisfaction and acceptance	28
Usefulness (self-reported)	22
Ease of use (self-reported)	14
Willingness to use	7
Number of websites or television advertisements with desired information	4
Technology literacy improvement	1

Discussion

The purpose of this study was to examine ways in which technology is being designed for historically underserved populations to facilitate or improve health care access and health outcomes. The reviewed studies focused on either (1) a defined historically underserved population, such as African Americans or people with a lower socioeconomic status, or (2) a historically underserved population, such as racial and ethnic minorities, as a group.

The results are organized into the four main questions. (1) Which types of technologies are used to address negative health outcomes for historically underserved populations? (2) At which disease, health problem, or potential problem is the technology aimed? (3) Which historically underserved groups are technology-based interventions designed for in the literature? (4) How are the health benefits and technologies evaluated?

Which Types of Technologies are Used to Address Negative Health Outcomes for Historically Underserved Populations?

The papers discussed 30 different types of technology; half (15) are typically used within a clinical setting, while the remaining 15 types are often used outside of a medical setting. Technologies that are often used outside of a clinical setting were mentioned in the majority of papers (102 papers) and included technologies such as videotapes, Internet, computer, and radio. While not originally created for the health care system, these types of technology were readily adapted to aid health consumers. If a historically underserved population is already familiar with and has access to this type of technology, the technology might be an appropriate platform choice. For instance, 34 papers used videos to relay health messages. Videos are readily understood and easily accessed by the majority of the US population and therefore likely a good choice for health

education or interventions aimed at historically underserved populations.

A total of 45 papers used technology typically used within a health care setting (eg, telemedicine, EHRs, or assisted reproductive technology). However, seven of these technologies (medical technology, electronic medical records, personal health records, high-technology hospitals, rapid HIV testing, implantable cardioverter defibrillator, and cochlear implants) were mentioned in only 1 paper [69,77,85,88,121,124,126]. Furthermore, telemedicine was the only type of health-specific technology mentioned in 10 or more papers.

Additionally, 16 papers discussed more than one type of technology, and the majority of these papers (14) mentioned two types of technology typically used outside of the medical office. The remaining 2 papers mentioned one type of each: one type of technology typically used at a clinic, and one type typically used outside of a clinic (telemedicine and videotapes [142], assistive technology and Internet [72]). None of the papers mentioned using more than one type of technology that is typically used inside a medical office.

Among the reviewed papers, videotapes were widely discussed as a method for interventions and educational tools (24 papers). Using videotapes instead of written materials to educate patients increased comprehension among breast cancer patients with low literacy skills [41]. Culturally tailored videotapes that employed characters of the same ethnic background as the patient influenced African American and Hispanic women on both a cognitive and emotional level [41]. Additionally, multiple studies showed increased trust among the patients when the narrator or main character of the educational videotape was the same ethnicity or race as the audience [47,144,150]. In addition, 2 studies demonstrated how storytelling can be used in videotapes to effectively communicate and educate patients about a specific health condition [78,122]. Videotapes were often complemented by other technology such as informational

brochures [63], the radio [125], the computer [45,47], self-efficacy and skill-building exercises [53], multimedia tools [102,136], and telemedicine [142].

The Internet is highly used by health care professionals for interventions and education. One study showed the increased benefit of the Internet to individuals with lower incomes and education levels despite their lower use of the Internet to access health information [151]. Women, minorities, and poverty-stricken individuals (who are also part of the population with the fastest-growing rate of HIV/acquired immunodeficiency syndrome [AIDS]) are those most likely to not have access to the Internet [59]. While many historically underserved populations have lower access to the Internet on a computer, they have higher usage rates of mobile Internet access on handheld devices [151]. Almost two-thirds of African American (64%) and Hispanic people (63%) have wireless access to the Internet. In fact, more African American and Hispanic people own cell phones (87%) than their white counterparts (80%) and, further, these historically underserved populations use their phone data functions more than their white counterparts do [9]. Gibbons suggested that, due to the high usage rates, these tools could improve patient engagement and be an effective mode for interventions [9]. Crilly and colleagues suggested wireless handheld devices as a viable alternative for patients who face barriers due to geography [151]. Eyrich-Garg conducted a study on homeless individuals who faced barriers due to geography. Of the participants in his study, nearly half (44%) owned a mobile phone [59]. Of this 44%, one-fifth had accessed the Internet via their mobile phone in the past 30 days. For this reason, Eyrich-Garg suggested that health care providers could disseminate health information to the homeless through use of mobile phones [59].

Using mobile phones as a means to send information via text messaging is mentioned in the literature as a viable option for racial groups. Similar to their usage of mobile phones, African Americans use text messaging more than their white counterparts do [55]. Samal et al found that text messaging was an acceptable mode of information and communication

technology for African American women in an urban sexually transmitted infection clinic [105]. Another study tested the feasibility of text messages as an HIV prevention method for young African American men. The results were positive and suggested that humor be used to initially engage the patient before providing an HIV fact later in the text [148]. While the research is new and applied to only a few select historically underserved populations, text messages are being used as a modality to disseminate health information to these populations.

To access the desired and undesired effects of technology and to search for relevant literature about a technology, a clear definition and delineation of technology is necessary [23]. When the technology is a surgical instrument or a piece of equipment, the definition is seldom a problem; however, other technologies are more complex and unformed, and require more thought to define (eg, wound care, fast-track surgery, or electronic medication). When a definition is created, the technology should be described from its material nature, its purpose, the degree of dissemination, and its maturity [23]. Kristensen and Sigmund suggested that the technology can be defined through a series of questions about how the technology is used for the disease or illness, or the technology [23]. To define the technologies in this review, we asked the following questions [23]. (1) Are there any special professional or technical requirements for operating the technology? (2) Are there factors that affect the application of the technology? (3) What is the purpose and application area of the technology? (4) At which disease, health problem, or potential problem is the technology aimed?

Questions 1 and 2 point to the need to effectively design technology that can overcome cultural differences that are exaggerated by the digital divide, health literacy, and language differences between historically underserved groups and the larger population. Every user needs to be able to operate and understand the technology to effectively access and use it to improve his or her health [23]. With regard to question 3, 23 identified 16 types of technology applications (Table 6). Question 4 is discussed in detail below.

Table 6. Types of technology

Application of technology	Number of papers
Intervention or education tool	62
Health management tool	19
Tool for communication with provider	6
Health record	5
Reproduction	5
Assistive technology	4
Information-gathering tool	3
Interpretation tool	2
Information and communication technology	1
Health information tool	1
Cardioverter defibrillation—medical technology	1
Cochlear implant—medical technology	1
Pharmacy tool	1
Drug advertisements	1
Knowledge acquisition	1
Health literacy assessment	1

At Which Disease, Health Problem, or Potential Problem is the Technology Aimed?

Although 23 health issues were discussed in the reviewed papers, general health was discussed in one-quarter of the papers (33). The next five most mentioned health issues (cancer, diabetes, HIV/AIDS, nutrition and physical activity, and sexually transmitted infections) were mentioned in a disproportionate number of papers (60), while the remaining 17 health issues were mentioned in only 37 papers. Furthermore, eight of the health issues (sensorineural hearing loss, organ donation, hepatitis C, health literacy, high blood pressure, poison control, hypertension, and child development) were mentioned a only single paper each.

Which Historically Underserved Groups are Technology-Based Interventions Designed for in the Literature?

The reviewed papers discussed 19 historically underserved groups. African American and Hispanic populations were mentioned at least twice as often (64 and 51 papers, respectively) as the second-largest target group (ie, women were mentioned in 26 papers). While African American and Hispanic populations were mentioned often, other racial and ethnic groups were rarely mentioned. Native Americans and Alaskan natives were mentioned in 4 papers and Asian Americans were mentioned in only 1 paper. The studies involving Native Americans and Alaskan natives provided an overview of the Indian Health Service [114] and evaluated the positive implementation of EHRs [115], a telehealth network [56], and library access through the Internet [146]. The studies demonstrated the importance of involving and empowering the community to successfully implement health IT [56,146]. More research is needed on this population to better understand the intricacies

of implementing health IT in the Indian Health Service. Only 1 paper mentioned Asian Americans [126]; however, this paper was not singularly about Asian Americans. The paper showed that white and Asian American children were more likely to receive cochlear implants than their Hispanic and African American counterparts [126].

While the majority of the papers did not mention gender, when gender was mentioned, women were discussed in 26 papers, while men were specifically discussed in only 4 papers. Of the 26 papers focused on women, 16 described health issues specific to women (7 papers on reproduction, 7 on breast cancer, and 2 on breast-feeding). The remaining 10 papers discussed health conditions that are not gender specific and that could affect males (3 papers on HIV, 2 on obesity, and 1 on the remaining health issues: general health, cardiovascular disease, sexually transmitted infections, nutrition, and cancer). Of the 4 papers dedicated to men, 1 discussed prostate cancer, which is specific only to men; however, the remaining 3 papers discussed HIV and sexually transmitted infections, which can also affect women. While it is understandable that papers discussing gender-specific health issues such as breast cancer or prostate cancer would focus on a single gender, 10 papers targeted only women and 3 papers targeted only men while addressing a non-gender-specific health issue.

It is important to note that attitudes toward technological interventions vary between historically underserved populations, not just between majority populations and historically underserved populations. A single intervention will not necessarily work for two separate racial or ethnic groups; interventions should be tailored to each population to be most effective. For instance, 1 study found that African American and Hispanic populations have different concerns regarding telemedicine [12]. While African American participants were

concerned by the physical absence of the health care professional and the ability to monitor their qualifications, Hispanic participants were concerned with whether telemedicine would be available to uninsured or undocumented individuals.

How are the Health Benefits and Technologies Evaluated?

The review papers used 12 types of evaluation. While we expected that most of the papers would use quantitative evaluation techniques, only half of the papers used these techniques. Objective evaluations were used in 90 papers (31 papers measured access or usage rates, 25 measured changes in health knowledge, 18 measured behavior changes, 10 measured biometric changes, 4 counted the number of websites or television advertisements with the desired information, 1 measured improvements in technology literacy, and 1 measured health-related quality of life). Self-reported measures were used in 93 papers (28 papers measured self-reported satisfaction or acceptance, 22 measured self-reported behavior changes, 22 measured self-reported usefulness, 14 measured self-reported ease of use, and 7 measure willingness to use the technology). Though 10 of the papers measured biometric changes, most of the papers did not evaluate the effects of the technology on health outcomes. Instead, the papers evaluated intermediate measures such as behavior changes or access rates of the technology.

Of the 67 papers that tested a culturally informed technology, 66 found the technology successful in at least one of the evaluated metrics; this points to the fact that health technology is an effective method to improve the health of historically underserved populations. The one study that did not have success aimed to improve HIV risk and sexual behaviors through a culturally appropriate educational video for 15- to 19-year-old black males [51]. Instead, the researchers suggested that an African American health educator conduct face-to-face interventions in order to have a greater impact.

Conclusion

This review illustrates how technology is being used by historically underserved populations to facilitate or improve their health care access and health and wellness outcomes. Synthesis of the literature points to the benefit of accounting for the end user's culture when designing health technology. A person's culture shapes how health information is received, what a health consumer considers a health problem, how symptoms are expressed, who should provide treatment, and what treatment should be provided [152]. The review conveys that culturally informed technology affects the health outcomes of the historically underserved populations facing health disparities in the United States.

Which Types of Technologies are Used to Address Negative Health Outcomes for Historically Underserved Populations?

The reviewed papers discussed 30 different types of technology, both those typically used inside a medical setting and those typically used outside of a medical setting. Health IT can lessen barriers facing historically underserved populations [11]. However, administrators and physicians should carefully analyze

the type of technology they choose to implement, as different types of technology are better than others at overcoming certain barriers. Since different historically underserved populations face distinct barriers, choosing a technology type should be an informed decision. For instance, people living in rural and underresourced areas face extra barriers related to provider availability and transportation [153]. The use of telemedicine, where the providers can be located in a different region, can overcome these barriers and aid historically underserved populations in accessing patient-centered care [4,37]. However, of the reviewed papers, only 4 discussed using telemedicine to aid underresourced or rural populations [4,12,37,35]. In another example, using culturally tailored technology that places little financial burden on the consumer and is easy to use, such as videotapes, television advertisements, and compact discs, can help mitigate health disparities facing individuals with low socioeconomic status [154]. This review provides evidence that these technologies have been implemented to help historically underserved populations (34 video papers, 8 television papers, 2 compact disc papers). Additionally, the type of technology with the greatest potential to aid individuals facing multiple chronic conditions is EHRs [155]. However, none of the 32 chronic disease papers (diabetes, HIV/AIDS, asthma, and hypertension) used EHRs.

Choosing an appropriate type of technology is not enough; the technology should be tailored toward the intended population, as personally relevant health technology is more likely than more broad technology to change behavior [1].

At Which Disease, Health Problem, or Potential Problem is the Technology Aimed?

The reviewed papers discussed 23 health issues, with 33 of the papers discussing general health concerns. Since the US federal government requires recipients of federal funds to provide language assistance services, including bilingual staff and interpreters, at no cost to the patient [152], it is surprising that health literacy was not mentioned in more papers. Technology can easily translate difficult health terms and issues into more easily understood concepts for laypeople. Recent reports by the Institute of Medicine and Agency for Healthcare Research and Quality (AHRQ) recommended that future research examine culture and cultural differences when measuring health literacy [156]. Specifically, the 2004 AHRQ report recommended that covariates such as socioeconomic status or education level should be further explored [157]. In addition, it is surprising that chronic diseases were not mentioned in more papers, as chronic diseases are the leading cause of health disparities [36]. Technology can help health consumers manage their overall health behaviors and medicine intake, and thus we expected that more papers would have discussed chronic diseases.

Nearly half of the papers discussed how the technology was used to create culturally informed interventions or educational tools. Obtaining access to culturally appropriate and accessible health education is a necessary piece of receiving high-quality, patient-centered care [154]. Similarly, the reviewed papers support Barrera et al's findings that culturally appropriate health interventions are more effective than usual care. However, there are important limitations to previous research [158]. Since

culturally adapted interventions are seldom directly compared with nonculturally informed interventions, it is difficult to state with certainty that the cultural aspect of the intervention was an important piece of the success of the intervention.

Which Historically Underserved Groups are Technology-Based Interventions Designed for in the Literature?

The papers included in this review highlight a relatively limited number of historically underserved groups (19). However, the review papers did discuss seven priority populations defined by the AHRQ. The AHRQ focused on seven priority populations as specified by Congress in the Healthcare Research and Quality Act of 1999: racial and ethnic minorities, low-income groups, women, children, older adults, residents of rural areas, and individuals with special health care needs, such as individuals with disabilities and individuals who need chronic care or end-of-life care [153]. According to AHRQ, racial minority groups are white people, black people, Asians, Native Hawaiian or other Pacific Islanders, American Indian and Alaska natives, and people who belong to more than one race; ethnic minority groups are Hispanic or Latino [153]. Within the reviewed papers, chronic care and disabilities were not discussed as characterizing historically underserved groups but were mentioned as health issues facing the different populations. In addition to the AHRQ priority populations, the reviewed papers discussed an additional four historically underserved groups: people who speak English as a second language, men, immigrants, and homeless people.

Of the 19 historically underserved groups discussed in the reviewed papers, 11 of these groups were discussed in fewer than 5 papers. Further, five groups (Asian Americans, immigrants, the homeless, people with AIDS, and people living with HIV) were discussed in a single paper. The discrepancy in the number of papers reviewed per historically underserved population is potentially problematic, as it can result in gaps in information regarding the less-studied populations [153]. Furthermore, understudied populations are left out of relevant discourse and in turn rendered invisible and powerless [159]. It is important to study all historically underserved groups to avoid this invisibility and bring awareness to the populations.

The reviewed papers tended to examine one identity that an individual might hold. In addition to studying historically underserved groups separately, researchers should examine populations from an intersectional theoretical perspective. Intersectionality refers to particular forms of intersecting oppressions [160] such as being both Hispanic and an older individual. The combination of these two identities creates different barriers for the patient than either single identity would create on its own. While the papers discussed 18 combinations of cultural groups, they did not adequately theorize the issue of intersectionality and therefore cannot fully understand the barriers and problems facing individuals within the group.

How are the Health Benefits and Technologies Evaluated?

We identified 2 main forms of evaluation in the reviewed papers: evaluation of health changes related to use of the technology and evaluation of the technology itself. A fraction of the papers

(23) used both types of evaluation. A wide range of evaluation metrics were used; about half of the papers (64) used self-reported measures as part of their evaluation, while 10 papers measured biometric changes. Even though previous research found self-reports to not be an accurate predictor of health information competencies [161], 32 papers used only self-reported metrics.

The reviewed papers did not include a validated method to evaluate the specific cultural aspect of the health technology. Design processes should be reported in the research so that best practices can be created for culturally relevant design methods. Only 13 of the reviewed papers provided detail on the design process of the interventions and educational tools. Future research should evaluate metrics for culturally informed health technology. These metrics will need to be adapted and changed for different cultural groups, as diverse cultural groups expect different criteria from their health technology.

Study Limitations

We followed systematic review methodology; however, this method has several limitations. Systematic reviews can only assess published work and report on the findings in those articles. Other potential limitations include the use of a single reviewer and the exclusion of studies regarding mental health, end-of-life care, and dental care.

Future Recommendations

More research about culturally informed technology for health is needed. In conjunction with this research, it is imperative for researchers to continue collecting data on cultural populations [162]. Gaps in knowledge about the access to and use of health services by historically underserved populations exist in terms of learning practices, methods to navigate services, and help-seeking behaviors [163]. Further research is necessary to understand the limitations of the data and avoid overgeneralizations [162]. Future recommendations include the following:

- Theoretical models and perspectives are needed to design culturally informed technologies.
- Methodologically, more research should be conducted to create a culturally informed approach to the design of health technology geared toward historically underserved populations. While methods should vary based on the technology, cultural population, and health issue, a broad methodology should be recommended for the future design of culturally informed health technology. This methodology might include formative research, which can aid researchers in overcoming their own implicit biases by using participatory methods to help them understand the population, create programs specific to the population's needs, and ensure the programs are acceptable to the population through pilot testing [164]. Formative research includes qualitative research methods such as focus groups, interviews with key informants, surveys, and field notes. When using formative research to develop culturally informed health IT, key informants might include cultural theorists.

- Financial incentives should be provided to organizations that adopt technology for historically underserved populations. The financial burden of purchasing, implementing, and maintaining health IT serves as a barrier to the adoption among underresourced providers who frequently serve lower socioeconomic patients [14].

Recommendations related to the type of technology chosen are as follows:

- When designing or implementing health technology for historically underserved populations, the type of technology should be carefully considered. Barriers to access and use of health IT differ between populations; different types of technology can be used to overcome distinct barriers. Therefore, the choice of technology type is important. Future research should create a comprehensive list of which types of technology would be most beneficial for each group. For instance, telemedicine is a useful tool to reach rural populations, and mobile telephones are a useful tool to reach African American populations.
- Trust and lack of cultural relevance have been found to be a barrier, as lack of trust in the technology, technical problems, or confusing instructions have a negative impact on adoption and usage rates among historically underserved populations [9]. More research is necessary to determine whether patients' culture changes their level of trust in culturally informed health IT.
- Future studies should examine how to best diffuse technology into a population [165]. When implementing an intervention, researchers should evaluate the readiness of the intended population.
- As new technology is invented and as the cost of current technology decreases, culturally informed health technology should be adapted. For instance, social media, which have

rapidly grown in the 21st century [166], should be further examined as a possible method to reach historically underserved populations. Social media have already started to enter the health care system through online patient communities such as PatientsLikeMe, QuitNet, and CureTogether. These networks create spaces for patients to discuss specific conditions and share their experiences. If access does not serve as a barrier, research shows that these social networks can be useful for historically underserved populations [9]. Social media may prove to be a cheaper way to access geographically isolated populations.

Recommendations related to the disease, health problem, or potential problem are the following:

- Future studies should use culturally informed health IT for chronic disease management. The emphasis on self-management support programs has shifted from pedagogical education with education content defined by health care professionals to an individualized approach that addresses the specific needs of a patient's situation [167]. Future research should examine how to best use technology to aid in the disease management of historically underserved populations with chronic diseases.

The recommendation related to the evaluation of the technology is the following:

- The evaluation process should be standardized to create a benchmark for culturally informed health IT. Participatory approaches should be used when possible to evaluate technologies, but metrics related to culturally informed design are needed. While research should dictate these metrics, possible metrics might include issues surrounding access, usability, perceived usefulness, and cultural appropriateness.

Acknowledgments

This publication is supported by grant 1UL1RR025011 from National Institutes of Health and the University of Wisconsin-Madison Systems Engineering Initiative for Patient Safety (<http://cqpi.engr.wisc.edu>). The Agency for Healthcare Research and Quality's training grant provided support for this project (<http://www.ahrq.gov/fund/training/rsrchtng.htm>).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Overview of reviewed papers.

[PDF File (Adobe PDF File), 161KB - [jmir_v14i3e78_app1.pdf](#)]

References

- Gibbons MC. A historical overview of health disparities and the potential of eHealth solutions. *J Med Internet Res* 2005;7(5):e50 [FREE Full text] [doi: [10.2196/jmir.7.5.e50](https://doi.org/10.2196/jmir.7.5.e50)] [Medline: [16403714](https://pubmed.ncbi.nlm.nih.gov/16403714/)]
- Pecukonis EV, Cornelius L, Parrish M. The future of health social work. *Soc Work Health Care* 2003;37(3):1-15. [doi: [10.1300/J010v37n03_01](https://doi.org/10.1300/J010v37n03_01)] [Medline: [14526873](https://pubmed.ncbi.nlm.nih.gov/14526873/)]
- Partida Y. Addressing language barriers: building response capacity for a changing nation. *J Gen Intern Med* 2007 Nov;22 Suppl 2:347-349. [doi: [10.1007/s11606-007-0367-1](https://doi.org/10.1007/s11606-007-0367-1)] [Medline: [17957423](https://pubmed.ncbi.nlm.nih.gov/17957423/)]
- Coen Buckwalter K, Lindsey Davis L, Wakefield BJ, Kienzle MG, Ann Murray M. Telehealth for elders and their caregivers in rural communities. *Fam Community Health* 2002 Oct;25(3):31-40. [Medline: [12802140](https://pubmed.ncbi.nlm.nih.gov/12802140/)]

5. Carter-Pokras O, Baquet C. What is a "health disparity"? *Public Health Rep* 2002;117(5):426-434. [Medline: [12500958](#)]
6. Betancourt JR, Green AR, Carrillo JE, Ananeh-Firempong O. Defining cultural competence: a practical framework for addressing racial/ethnic disparities in health and health care. *Public Health Rep* 2003;118(4):293-302. [Medline: [12815076](#)]
7. Baquet CR, Carter-Pokras O, Bengen-Seltzer B. Healthcare disparities and models for change. *Am J Manag Care* 2004 Sep;10 Spec No:SP5-S11 [[FREE Full text](#)] [Medline: [15481431](#)]
8. US Centers for Disease Control and Prevention, National Center for Health Statistics. 2011. Final Review: Healthy People 2010 Executive Summary URL: http://www.cdc.gov/nchs/data/hpdata2010/hpdata2010_final_review_executive_summary.pdf [accessed 2012-05-22] [[WebCite Cache ID 67rGLcYDx](#)]
9. Christopher Gibbons M. Use of health information technology among racial and ethnic underserved communities. *Perspect Health Inf Manag* 2011 Jan;8:1f. [Medline: [21307989](#)]
10. World Health Organization. 2009. Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth URL: http://www.who.int/goe/publications/goe_telemedicine_2010.pdf [accessed 2012-05-24] [[WebCite Cache ID 67uFzmAKU](#)]
11. Leavitt MO. Health Information Technology Initiative: Major Accomplishments: 2004-2006. Washington, DC: US Department of Health and Human Services; 2011.
12. George SM, Hamilton A, Baker R. Pre-experience perceptions about telemedicine among African Americans and Latinos in South Central Los Angeles. *Telemed J E Health* 2009;15(6):525-530. [doi: [10.1089/tmj.2009.0152](#)] [Medline: [19566397](#)]
13. Fiscella K, Geiger HJ. Health information technology and quality improvement for community health centers. *Health Aff (Millwood)* 2006;25(2):405-412. [doi: [10.1377/hlthaff.25.2.405](#)] [Medline: [16522580](#)]
14. Blumenthal D, DesRoches C, Foubister V. George Washington University, Massachusetts General Hospital, Robert Wood Johnson Foundation. 2008. Health Information Technology in the United States?: Where We Stand, 2008 URL: <http://www.rwjf.org/files/research/3297.31831.hitreport.pdf> [accessed 2012-05-22] [[WebCite Cache ID 67rGtVzv6](#)]
15. Millery M, Kukafka R. Health information technology and quality of health care: strategies for reducing disparities in underresourced settings. *Med Care Res Rev* 2010 Oct;67(5 Suppl):268S-298S. [doi: [10.1177/1077558710373769](#)] [Medline: [20675348](#)]
16. Li C, West-Strum D. Patient panel of underserved populations and adoption of electronic medical record systems by office-based physicians. *Health Serv Res* 2010 Aug;45(4):963-984. [doi: [10.1111/j.1475-6773.2010.01113.x](#)] [Medline: [20403062](#)]
17. Custodio R, Gard AM, Graham G. Health information technology: addressing health disparity by improving quality, increasing access, and developing workforce. *J Health Care Poor Underserved* 2009 May;20(2):301-307. [doi: [10.1353/hpu.0.0142](#)] [Medline: [19395831](#)]
18. Cooper H. *Synthesizing Research: A Guide for Literature Reviews*. 3rd edition. Thousand Oaks, CA: Sage Publications; 1998:86-88.
19. Paniagua FA. *Assessing and Treating Culturally Diverse Clients: A Practical Guide*. 3rd edition. Thousand Oaks, CA: Sage Publications; 2005:1-4.
20. 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](#)]
21. Larson JS. The conceptualization of health. *Med Care Res Rev* 1999 Jun;56(2):123-136. [Medline: [10373720](#)]
22. Varshney U. Pervasive healthcare and wireless health monitoring. *Mobile Netw Appl* 2007 Jul;12(2-3):113-127 [[FREE Full text](#)]
23. Kristensen F, Sigmund H. *Health Technology Assessment Handbook*. 2nd edition. Copenhagen: Danish Centre for Health Technology Assessment, National Board of Health; 2007. URL: http://www.sst.dk/publ/Publ2008/MTV/Metode/HTA_Handbook_net_final.pdf [accessed 2012-05-24] [[WebCite Cache ID 67uGdn71y](#)]
24. Thompson TG, Brailer DJ. *The Decade of Health Information Technology: Delivering Consumer-Centric and Information-Rich Health Care*. Washington, DC: US Department of Health and Human Services, Office of the National Coordinator for Health Information Technology; 2004. URL: http://www.providersedge.com/ehdocs/ehr_articles/The_Decade_of_HIT-Delivering_Customer-centric_and_Info-rich_HC.pdf [accessed 2012-05-24] [[WebCite Cache ID 67uGm2BqF](#)]
25. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc* 2009;16(4):550-560 [[FREE Full text](#)] [doi: [10.1197/jamia.M2888](#)] [Medline: [19390112](#)]
26. Magro A, Swarz J, Ousley A. CancerSPACE: an interactive e-learning tool aimed to improve cancer screening rates. *J Comput Mediat Commun* 2010;15(3):482-499. [doi: [10.1111/j.1083-6101.2010.01512.x](#)]
27. World Health Organization Western Pacific Region. 2006. *Electronic Health Records: A Manual for Developing Countries* URL: <http://www.wpro.who.int/publications/docs/EHRmanual.pdf> [accessed 2012-05-24] [[WebCite Cache ID 67uGoCk49](#)]
28. Garets D, Davis M. *Healthcare Informatics Online*. 2005 Oct. *Electronic Patient Records: EMRs and EHRs* URL: http://www.providersedge.com/ehdocs/ehr_articles/Electronic_Patient_Records-EMRs_and_EHRs.pdf [accessed 2012-05-24] [[WebCite Cache ID 67uGqG2mB](#)]

29. Markle Foundation. 2006 Dec 7. Connecting Americans to Their Health Care: A Common Framework for Networked Personal Health Information URL: <http://www.accessmyhealth.org/documents/A%20Common%20Framework%20Networked%20PHRs.pdf> [accessed 2012-05-24] [WebCite Cache ID 67uGr0DIQ]
30. Grinter RE, Siek KA, Grimes A. Is wellness informatics a field of human-centered health informatics? *interactions* 2010;17(1):76-79. [doi: [10.1145/1649475.1649494](https://doi.org/10.1145/1649475.1649494)]
31. Van De Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. *J Med Internet Res* 2010;12(2):e18 [FREE Full text] [doi: [10.2196/jmir.1350](https://doi.org/10.2196/jmir.1350)] [Medline: [20542857](https://pubmed.ncbi.nlm.nih.gov/20542857/)]
32. Abbatangelo-Gray J, Byrd-Bredbenner C, Austin SB. Health and nutrient content claims in food advertisements on Hispanic and mainstream prime-time television. *J Nutr Educ Behav* 2008;40(6):348-354. [doi: [10.1016/j.jneb.2008.01.003](https://doi.org/10.1016/j.jneb.2008.01.003)] [Medline: [18984490](https://pubmed.ncbi.nlm.nih.gov/18984490/)]
33. Alvaro EM, Siegel JT, Crano WD, Dominick A. A mass mediated intervention on Hispanic live kidney donation. *J Health Commun* 2010 Jun;15(4):374-387. [doi: [10.1080/10810731003753133](https://doi.org/10.1080/10810731003753133)] [Medline: [20574876](https://pubmed.ncbi.nlm.nih.gov/20574876/)]
34. Anderson RM, Funnell MM, Arnold MS, Barr PA, Edwards GJ, Fitzgerald JT. Assessing the cultural relevance of an education program for urban African Americans with diabetes. *Diabetes Educ* 2000;26(2):280-289. [Medline: [10865593](https://pubmed.ncbi.nlm.nih.gov/10865593/)]
35. Arora S, Thornton K, Jenkusky SM, Parish B, Scaletti JV. Project ECHO: linking university specialists with rural and prison-based clinicians to improve care for people with chronic hepatitis C in New Mexico. *Public Health Rep* 2007 Jan;122 Suppl 2:74-77. [Medline: [17542458](https://pubmed.ncbi.nlm.nih.gov/17542458/)]
36. Baig AA, Wilkes AE, Davis AM, Peek ME, Huang ES, Bell DS, et al. The use of quality improvement and health information technology approaches to improve diabetes outcomes in African American and Hispanic patients. *Med Care Res Rev* 2010 Oct;67(5 Suppl):163S-197S. [doi: [10.1177/1077558710374621](https://doi.org/10.1177/1077558710374621)] [Medline: [20675350](https://pubmed.ncbi.nlm.nih.gov/20675350/)]
37. Balamurugan A, Hall-Barrow J, Blevins MA, Brech D, Phillips M, Holley E, et al. A pilot study of diabetes education via telemedicine in a rural underserved community--opportunities and challenges: a continuous quality improvement process. *Diabetes Educ* 2009;35(1):147-154. [doi: [10.1177/0145721708326988](https://doi.org/10.1177/0145721708326988)] [Medline: [19244570](https://pubmed.ncbi.nlm.nih.gov/19244570/)]
38. Bell DS, Daly DM, Robinson P. Is there a digital divide among physicians? A geographic analysis of information technology in Southern California physician offices. *J Am Med Inform Assoc* 2003;10(5):484-493 [FREE Full text] [doi: [10.1197/jamia.M1341](https://doi.org/10.1197/jamia.M1341)] [Medline: [12807804](https://pubmed.ncbi.nlm.nih.gov/12807804/)]
39. Bertera EM, Tran BQ, Wuertz EM, Bonner A. A study of the receptivity to telecare technology in a community-based elderly minority population. *J Telemed Telecare* 2007 Jan;13(7):327-332. [doi: [10.1258/135763307782215325](https://doi.org/10.1258/135763307782215325)] [Medline: [17958932](https://pubmed.ncbi.nlm.nih.gov/17958932/)]
40. Black MM, Teti LO. Promoting mealtime communication between adolescent mothers and their infants through videotape. *Pediatrics* 1997 Mar;99(3):432-437. [Medline: [9041301](https://pubmed.ncbi.nlm.nih.gov/9041301/)]
41. Borrayo EA. Where's Maria? A video to increase awareness about breast cancer and mammography screening among low-literacy Latinas. *Prev Med* 2004 Jul;39(1):99-110. [doi: [10.1016/j.ypmed.2004.03.024](https://doi.org/10.1016/j.ypmed.2004.03.024)] [Medline: [15207991](https://pubmed.ncbi.nlm.nih.gov/15207991/)]
42. Buchanan LM, Khazanchi D. A PDA intervention to sustain smoking cessation in clients with socioeconomic vulnerability. *West J Nurs Res* 2010 Apr;32(3):281-304. [doi: [10.1177/0193945909352664](https://doi.org/10.1177/0193945909352664)] [Medline: [20581398](https://pubmed.ncbi.nlm.nih.gov/20581398/)]
43. Calderon Y, Leider J, Hailpern S, Haughey M, Ghosh R, Lombardi P, et al. A randomized control trial evaluating the educational effectiveness of a rapid HIV posttest counseling video. *Sex Transm Dis* 2009 Apr;36(4):207-210. [doi: [10.1097/OLQ.0b013e318191ba3f](https://doi.org/10.1097/OLQ.0b013e318191ba3f)] [Medline: [19265735](https://pubmed.ncbi.nlm.nih.gov/19265735/)]
44. Calles-Escandón J, Hunter JC, Langdon SE, Gómez EM, Duren-Winfield VT, Woods KF. La Clínica del Pueblo: a model of collaboration between a private media broadcasting corporation and an academic medical center for health education for North Carolina Latinos. *J Immigr Minor Health* 2009 Dec;11(6):513-519. [doi: [10.1007/s10903-008-9223-6](https://doi.org/10.1007/s10903-008-9223-6)] [Medline: [19156524](https://pubmed.ncbi.nlm.nih.gov/19156524/)]
45. Campbell MK, Honess-Morreale L, Farrell D, Carbone E, Brasure M. A tailored multimedia nutrition education pilot program for low-income women receiving food assistance. *Health Educ Res* 1999 Apr;14(2):257-267 [FREE Full text] [Medline: [10387505](https://pubmed.ncbi.nlm.nih.gov/10387505/)]
46. Carter EL, Nunlee-Bland G, Callender C. A patient-centric, provider-assisted diabetes telehealth self-management intervention for urban minorities. *Perspect Health Inf Manag* 2011 Jan;8:1b. [Medline: [21307985](https://pubmed.ncbi.nlm.nih.gov/21307985/)]
47. Champion VL, Springston JK, Zollinger TW, Saywell RM, Monahan PO, Zhao Q, et al. Comparison of three interventions to increase mammography screening in low income African American women. *Cancer Detect Prev* 2006 Jan;30(6):535-544. [doi: [10.1016/j.cdp.2006.10.003](https://doi.org/10.1016/j.cdp.2006.10.003)] [Medline: [17110056](https://pubmed.ncbi.nlm.nih.gov/17110056/)]
48. Chan YF, Alagappan K, Rella J, Bentley S, Soto-Greene M, Martin M. Interpreter services in emergency medicine. *J Emerg Med* 2010 Feb;38(2):133-139. [doi: [10.1016/j.jemermed.2007.09.045](https://doi.org/10.1016/j.jemermed.2007.09.045)] [Medline: [18571358](https://pubmed.ncbi.nlm.nih.gov/18571358/)]
49. Choi N. Relationship between health service use and health information technology use among older adults: analysis of the US National Health Interview Survey. *J Med Internet Res* 2011 Jan;13(2):e33 [FREE Full text] [doi: [10.2196/jmir.1753](https://doi.org/10.2196/jmir.1753)] [Medline: [21752784](https://pubmed.ncbi.nlm.nih.gov/21752784/)]
50. Clifton GD, Byer H, Heaton K, Haberman DJ, Gill H. Provision of pharmacy services to underserved populations via remote dispensing and two-way videoconferencing. *Am J Health Syst Pharm* 2003 Dec 15;60(24):2577-2582. [Medline: [14750497](https://pubmed.ncbi.nlm.nih.gov/14750497/)]

51. DeLamater J, Wagstaff DA, Havens KK. The impact of a culturally appropriate STD/AIDS education intervention on black male adolescents' sexual and condom use behavior. *Health Educ Behav* 2000 Aug;27(4):454-470. [Medline: [10929753](#)]
52. Delorme DE, Huh J, Reid LN. Evaluation, use, and usefulness of prescription drug information sources among Anglo and Hispanic Americans. *J Health Commun* 2010 Jan;15(1):18-38. [doi: [10.1080/10810730903460526](#)] [Medline: [20390975](#)]
53. DeMarco RF, Kendricks M, Dolmo Y, Looby SE, Rinne K. The effect of prevention messages and self-efficacy skill building with inner-city women at risk for HIV infection. *J Assoc Nurses AIDS Care* 2009;20(4):283-292. [doi: [10.1016/j.jana.2009.02.007](#)] [Medline: [19576545](#)]
54. DeMarco R, Norris AE. Culturally relevant HIV interventions: transcending ethnicity. *J Cult Divers* 2004 Jan;11(2):65-68. [Medline: [15453006](#)]
55. Denizard-Thompson NM, Feiereisel KB, Stevens SF, Miller DP, Wofford JL. The digital divide at an urban community health center: implications for quality improvement and health care access. *J Community Health* 2011 Jun;36(3):456-460. [doi: [10.1007/s10900-010-9327-5](#)] [Medline: [21086028](#)]
56. Doorenbos AZ, Demiris G, Towle C, Kundu A, Revels L, Colven R, et al. Developing the Native People for Cancer Control Telehealth Network. *Telemed J E Health* 2011;17(1):30-34. [doi: [10.1089/tmj.2010.0101](#)] [Medline: [21214371](#)]
57. Eddens KS, Kreuter MW, Morgan JC, Beatty KE, Jasim SA, Garibay L, et al. Disparities by race and ethnicity in cancer survivor stories available on the web. *J Med Internet Res* 2009;11(4):e50 [FREE Full text] [doi: [10.2196/jmir.1163](#)] [Medline: [19945948](#)]
58. Essien EJ, Meshack AF, Peters RJ, Ogungbade GO, Osemene NI. Strategies to prevent HIV transmission among heterosexual African-American men. *BMC Public Health* 2005 Jan 7;5:3 [FREE Full text] [doi: [10.1186/1471-2458-5-3](#)] [Medline: [15638937](#)]
59. Eyrich-Garg KM. Mobile phone technology: a new paradigm for the prevention, treatment, and research of the non-sheltered "street" homeless? *J Urban Health* 2010 May;87(3):365-380. [doi: [10.1007/s11524-010-9456-2](#)] [Medline: [20397058](#)]
60. Feinberg EC, Larsen FW, Catherino WH, Zhang J, Armstrong AY. Comparison of assisted reproductive technology utilization and outcomes between Caucasian and African American patients in an equal-access-to-care setting. *Fertil Steril* 2006 Apr;85(4):888-894. [doi: [10.1016/j.fertnstert.2005.10.028](#)] [Medline: [16580370](#)]
61. Finkelstein J, Cha E, Dennison CR. Exploring feasibility of home telemanagement in African Americans with congestive heart failure. *Stud Health Technol Inform* 2010 Jan;160(Pt 1):535-539. [Medline: [20841744](#)]
62. Fitzner K, Dietz DA, Moy E. How innovative treatment models and data use are improving diabetes care among older African American adults. *Popul Health Manag* 2011 Jun;14(3):143-155. [doi: [10.1089/pop.2010.0012](#)] [Medline: [21323464](#)]
63. Frosch DL, Légaré F, Mangione CM. Using decision aids in community-based primary care: a theory-driven evaluation with ethnically diverse patients. *Patient Educ Couns* 2008 Dec;73(3):490-496. [doi: [10.1016/j.pec.2008.07.040](#)] [Medline: [18771875](#)]
64. Gans KM, Kumanyika SK, Lovell HJ, Risica PM, Goldman R, Odoms-Young A, et al. The development of SisterTalk: a cable TV-delivered weight control program for black women. *Prev Med* 2003 Dec;37(6 Pt 1):654-667. [Medline: [14636799](#)]
65. Garcia-Castillo D, Feters MD. Quality in medical translations: a review. *J Health Care Poor Underserved* 2007 Feb;18(1):74-84. [doi: [10.1353/hpu.2007.0009](#)] [Medline: [17337799](#)]
66. Gilbert MR, Masucci M, Homko C, Bove AA. Theorizing the digital divide: information and communication technology use frameworks among poor women using a telemedicine system. *Geoforum* 2008;39(2):438-425. [doi: [10.1016/j.geoforum.2007.09.001](#)]
67. Gordon NP, Iribarren C. Health-related characteristics and preferred methods of receiving health education according to dominant language among Latinos aged 25 to 64 in a large Northern California health plan. *BMC Public Health* 2008 Jan;8:305 [FREE Full text] [doi: [10.1186/1471-2458-8-305](#)] [Medline: [18782454](#)]
68. Grindel CG, Brown L, Caplan L, Blumenthal D. The effect of breast cancer screening messages on knowledge, attitudes, perceived risk, and mammography screening of African American women in the rural South. *Oncol Nurs Forum* 2004 Jul;31(4):801-808. [doi: [10.1188/04.ONF.801-808](#)] [Medline: [15252434](#)]
69. Groeneveld PW, Sonnad SS, Lee AK, Asch DA, Shea JE. Racial differences in attitudes toward innovative medical technology. *J Gen Intern Med* 2006 Jun;21(6):559-563. [doi: [10.1111/j.1525-1497.2006.00453.x](#)] [Medline: [16808736](#)]
70. Gross SM, Caulfield LE, Bentley ME, Bronner Y, Kessler L, Jensen J, et al. Counseling and motivational videotapes increase duration of breast-feeding in African-American WIC participants who initiate breast-feeding. *J Am Diet Assoc* 1998 Feb;98(2):143-148. [Medline: [12515413](#)]
71. Guzman A. Learning about assistive technology: Hispanics and a national sample. *Assist Technol* 2008 Jan;20(1):48-54. [doi: [10.1080/10400435.2008.10131931](#)] [Medline: [18751579](#)]
72. Guzman A, Ostrander N. Hispanics' awareness of assistive technology. *Assist Technol* 2009 Jan;21(1):28-34. [doi: [10.1080/10400430902945868](#)] [Medline: [19719061](#)]
73. Helft PR, Eckles RE, Johnson-Calley CS, Daugherty CK. Use of the internet to obtain cancer information among cancer patients at an urban county hospital. *J Clin Oncol* 2005 Aug 1;23(22):4954-4962. [doi: [10.1200/JCO.2005.09.621](#)] [Medline: [16051948](#)]

74. Helion AM, Reddy DM, Kies AL, Morris DR, Wilson CM. Influence of communicator's race on efficacy of an HIV/STD prevention intervention among African American and Caucasian college students. *Public Health Nurs* 2008;25(5):440-450. [doi: [10.1111/j.1525-1446.2008.00728.x](https://doi.org/10.1111/j.1525-1446.2008.00728.x)] [Medline: [18816361](https://pubmed.ncbi.nlm.nih.gov/18816361/)]
75. Henderson VR, Kelly B. Food advertising in the age of obesity: content analysis of food advertising on general market and african american television. *J Nutr Educ Behav* 2005;37(4):191-196. [Medline: [16029689](https://pubmed.ncbi.nlm.nih.gov/16029689/)]
76. Hoffman AM, Redding CA, Goldberg D, Añel D, Prochaska JO, Meyer PM, et al. Computer expert systems for African-American smokers in physicians offices: a feasibility study. *Prev Med* 2006 Sep;43(3):204-211. [doi: [10.1016/j.ypmed.2006.03.025](https://doi.org/10.1016/j.ypmed.2006.03.025)] [Medline: [16780939](https://pubmed.ncbi.nlm.nih.gov/16780939/)]
77. Horan TA, Botts NE, Burkhard RJ. A multidimensional view of personal health systems for underserved populations. *J Med Internet Res* 2010 Jan;12(3):e32 [FREE Full text] [doi: [10.2196/jmir.1355](https://doi.org/10.2196/jmir.1355)] [Medline: [20685644](https://pubmed.ncbi.nlm.nih.gov/20685644/)]
78. Houston TK, Allison JJ, Sussman M, Horn W, Holt CL, Trobaugh J, et al. Culturally appropriate storytelling to improve blood pressure: a randomized trial. *Ann Intern Med* 2011 Jan;154(2):77-84. [Medline: [21242364](https://pubmed.ncbi.nlm.nih.gov/21242364/)]
79. Jantz C, Anderson J, Gould SM. Using computer-based assessments to evaluate interactive multimedia nutrition education among low-income predominantly Hispanic participants. *J Nutr Educ Behav* 2002;34(5):252-260. [Medline: [12559060](https://pubmed.ncbi.nlm.nih.gov/12559060/)]
80. Jha AK, Bates DW, Jenter C, Orav EJ, Zheng J, Cleary P, et al. Electronic health records: use, barriers and satisfaction among physicians who care for black and Hispanic patients. *J Eval Clin Pract* 2009 Feb;15(1):158-163. [doi: [10.1111/j.1365-2753.2008.00975.x](https://doi.org/10.1111/j.1365-2753.2008.00975.x)] [Medline: [18759752](https://pubmed.ncbi.nlm.nih.gov/18759752/)]
81. Kalichman SC, Cherry C, Browne-Sperling F. Effectiveness of a video-based motivational skills-building HIV risk-reduction intervention for inner-city African American men. *J Consult Clin Psychol* 1999 Dec;67(6):959-966. [Medline: [10596517](https://pubmed.ncbi.nlm.nih.gov/10596517/)]
82. Kalichman SC, Weinhardt L, Benotsch E, Cherry C. Closing the digital divide in HIV/AIDS care: development of a theory-based intervention to increase Internet access. *AIDS Care* 2002 Aug;14(4):523-537. [doi: [10.1080/09540120208629670](https://doi.org/10.1080/09540120208629670)] [Medline: [12204154](https://pubmed.ncbi.nlm.nih.gov/12204154/)]
83. Kaye HS, Yeager P, Reed M. Disparities in usage of assistive technology among people with disabilities. *Assist Technol* 2008 Jan;20(4):194-203. [doi: [10.1080/10400435.2008.10131946](https://doi.org/10.1080/10400435.2008.10131946)] [Medline: [19160906](https://pubmed.ncbi.nlm.nih.gov/19160906/)]
84. Kelly NR, Huffman LC, Mendoza FS, Robinson TN. Effects of a videotape to increase use of poison control centers by low-income and Spanish-speaking families: a randomized, controlled trial. *Pediatrics* 2003 Jan;111(1):21-26. [Medline: [12509549](https://pubmed.ncbi.nlm.nih.gov/12509549/)]
85. Kim TH, Samson LF, Lu N. Racial/ethnic disparities in the utilization of high-technology hospitals. *J Natl Med Assoc* 2010 Sep;102(9):803-810. [Medline: [20922924](https://pubmed.ncbi.nlm.nih.gov/20922924/)]
86. Kreuter MW, Holmes K, Alcaraz K, Kalesan B, Rath S, Richert M, et al. Comparing narrative and informational videos to increase mammography in low-income African American women. *Patient Educ Couns* 2010 Dec;81 Suppl:S6-14. [doi: [10.1016/j.pec.2010.09.008](https://doi.org/10.1016/j.pec.2010.09.008)] [Medline: [21071167](https://pubmed.ncbi.nlm.nih.gov/21071167/)]
87. Leeman-Castillo B, Beaty B, Raghunath S, Steiner J, Bull S. LUCAR: using computer technology to battle heart disease among Latinos. *Am J Public Health* 2010 Feb;100(2):272-275. [doi: [10.2105/AJPH.2009.162115](https://doi.org/10.2105/AJPH.2009.162115)] [Medline: [20019305](https://pubmed.ncbi.nlm.nih.gov/20019305/)]
88. Li J, Garcia S, Castro HK, DeForge BR, Hise MK, Finkelstein J. Acceptance and expectations of information technology to support hypertension self-care in African Americans: a qualitative inquiry. *AMIA Annu Symp Proc* 2007:1032. [Medline: [18694130](https://pubmed.ncbi.nlm.nih.gov/18694130/)]
89. 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](https://pubmed.ncbi.nlm.nih.gov/16978003/)]
90. Mandl KD, Katz SB, Kohane IS. Social equity and access to the World Wide Web and E-mail: implications for design and implementation of medical applications. *Proc AMIA Symp* 1998:215-219. [Medline: [9929213](https://pubmed.ncbi.nlm.nih.gov/9929213/)]
91. Mackert M, Kahlor L, Tyler D, Gustafson J. Designing e-health interventions for low-health-literate culturally diverse parents: addressing the obesity epidemic. *Telemed J E Health* 2009 Sep;15(7):672-677. [doi: [10.1089/tmj.2009.0012](https://doi.org/10.1089/tmj.2009.0012)] [Medline: [19694596](https://pubmed.ncbi.nlm.nih.gov/19694596/)]
92. McCarthy-Keith DM, Schisterman EF, Robinson RD, O'Leary K, Lucidi RS, Armstrong AY. Will decreasing assisted reproduction technology costs improve utilization and outcomes among minority women? *Fertil Steril* 2010 Dec;94(7):2587-2589. [doi: [10.1016/j.fertnstert.2010.02.021](https://doi.org/10.1016/j.fertnstert.2010.02.021)] [Medline: [20356585](https://pubmed.ncbi.nlm.nih.gov/20356585/)]
93. Mendelsohn AL, Valdez PT, Flynn V, Foley GM, Berkule SB, Tomopoulos S, et al. Use of videotaped interactions during pediatric well-child care: impact at 33 months on parenting and on child development. *J Dev Behav Pediatr* 2007 Jun;28(3):206-212. [doi: [10.1097/DBP.0b013e3180324d87](https://doi.org/10.1097/DBP.0b013e3180324d87)] [Medline: [17565287](https://pubmed.ncbi.nlm.nih.gov/17565287/)]
94. 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](https://doi.org/10.1258/135763307781644915)] [Medline: [17785031](https://pubmed.ncbi.nlm.nih.gov/17785031/)]
95. Mosnaim GS, Cohen MS, Rhoads CH, Rittner SS, Powell LH. Use of MP3 players to increase asthma knowledge in inner-city African-American adolescents. *Int J Behav Med* 2008 Jan;15(4):341-346. [doi: [10.1080/10705500802365656](https://doi.org/10.1080/10705500802365656)] [Medline: [19005935](https://pubmed.ncbi.nlm.nih.gov/19005935/)]
96. Outley CW, Taddese A. A content analysis of health and physical activity messages marketed to African American children during after-school television programming. *Arch Pediatr Adolesc Med* 2006 Apr;160(4):432-435. [doi: [10.1001/archpedi.160.4.432](https://doi.org/10.1001/archpedi.160.4.432)] [Medline: [16585490](https://pubmed.ncbi.nlm.nih.gov/16585490/)]

97. Nicklas TA, Goh ET, Goodell LS, Acuff DS, Reiher R, Buday R, et al. Impact of commercials on food preferences of low-income, minority preschoolers. *J Nutr Educ Behav* 2011;43(1):35-41. [doi: [10.1016/j.jneb.2009.11.007](https://doi.org/10.1016/j.jneb.2009.11.007)] [Medline: [20851053](https://pubmed.ncbi.nlm.nih.gov/20851053/)]
98. Noell J, Ary D, Duncan T. Development and evaluation of a sexual decision-making and social skills program: "the choice is yours--preventing HIV/STDs". *Health Educ Behav* 1997 Feb;24(1):87-101. [Medline: [9112100](https://pubmed.ncbi.nlm.nih.gov/9112100/)]
99. Olshefsky AM, Zive MM, Scolari R, Zuñiga M. Promoting HIV risk awareness and testing in Latinos living on the U.S.-Mexico border: the Tú No Me Conoces social marketing campaign. *AIDS Educ Prev* 2007 Oct;19(5):422-435. [doi: [10.1521/aeap.2007.19.5.422](https://doi.org/10.1521/aeap.2007.19.5.422)] [Medline: [17967112](https://pubmed.ncbi.nlm.nih.gov/17967112/)]
100. Pérez-Escamilla R, Himmelgreen D, Bonello H, Peng YK, Mengual G, González A, et al. Marketing nutrition among urban Latinos: the SALUD! campaign. *J Am Diet Assoc* 2000 Jun;100(6):698-701. [Medline: [10863575](https://pubmed.ncbi.nlm.nih.gov/10863575/)]
101. Powe BD, Weinrich S. An intervention to decrease cancer fatalism among rural elders. *Oncol Nurs Forum* 1999 Apr;26(3):583-588. [Medline: [10214599](https://pubmed.ncbi.nlm.nih.gov/10214599/)]
102. Prokhorov AV, Kelder SH, Shegog R, Murray N, Peters R, Agurcia-Parker C, et al. Impact of A Smoking Prevention Interactive Experience (ASPIRE), an interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high-school students. *Nicotine Tob Res* 2008 Sep;10(9):1477-1485. [doi: [10.1080/14622200802323183](https://doi.org/10.1080/14622200802323183)] [Medline: [19023839](https://pubmed.ncbi.nlm.nih.gov/19023839/)]
103. Ralph LJ, Berglas NF, Schwartz SL, Brindis CD. Finding teens in TheirSpace: using social networking sites to connect youth to sexual health services. *Sex Res Social Policy* 2011;8(1):38-49. [doi: [10.1007/s13178-001-0043-4](https://doi.org/10.1007/s13178-001-0043-4)]
104. Riegel B, Carlson B, Glaser D, Romero T. Randomized controlled trial of telephone case management in Hispanics of Mexican origin with heart failure. *J Card Fail* 2006 Apr;12(3):211-219. [doi: [10.1016/j.cardfail.2006.01.005](https://doi.org/10.1016/j.cardfail.2006.01.005)] [Medline: [16624687](https://pubmed.ncbi.nlm.nih.gov/16624687/)]
105. Samal L, Hutton HE, Erbeling EJ, Brandon ES, Finkelstein J, Chander G. Digital divide: variation in internet and cellular phone use among women attending an urban sexually transmitted infections clinic. *J Urban Health* 2010 Jan;87(1):122-128. [doi: [10.1007/s11524-009-9415-y](https://doi.org/10.1007/s11524-009-9415-y)] [Medline: [19941085](https://pubmed.ncbi.nlm.nih.gov/19941085/)]
106. Sánchez JP, Kaltwasser S, McClellan M, Burton WB, Blank A, Calderon Y. Educational video tool to increase syphilis knowledge among black and Hispanic male patients. *J Health Care Poor Underserved* 2010 Feb;21(1):371-385. [doi: [10.1353/hpu.0.0248](https://doi.org/10.1353/hpu.0.0248)] [Medline: [20173276](https://pubmed.ncbi.nlm.nih.gov/20173276/)]
107. Sánchez JP, Guiliames C, Sánchez NF, Calderon Y, Burton WB. Video tool to promote knowledge of syphilis among black and Hispanic men recruited from clinical and non-clinical settings. *J Community Health* 2010 Jun;35(3):220-228. [doi: [10.1007/s10900-010-9239-4](https://doi.org/10.1007/s10900-010-9239-4)] [Medline: [20151183](https://pubmed.ncbi.nlm.nih.gov/20151183/)]
108. Scheinmann R, Chiasson MA, Hartel D, Rosenberg TJ. Evaluating a bilingual video to improve infant feeding knowledge and behavior among immigrant Latina mothers. *J Community Health* 2010 Oct;35(5):464-470. [doi: [10.1007/s10900-009-9202-4](https://doi.org/10.1007/s10900-009-9202-4)] [Medline: [20039195](https://pubmed.ncbi.nlm.nih.gov/20039195/)]
109. Schillinger D, Hammer H, Wang F, Palacios J, McLean I, Tang A, et al. Seeing in 3-D: examining the reach of diabetes self-management support strategies in a public health care system. *Health Educ Behav* 2008 Oct;35(5):664-682. [doi: [10.1177/1090198106296772](https://doi.org/10.1177/1090198106296772)] [Medline: [17513690](https://pubmed.ncbi.nlm.nih.gov/17513690/)]
110. Schitai A. Caring for Hispanic patients interactively: simulations and practices for allied health professionals. *J Nurses Staff Dev* 2004;20(1):50-55. [Medline: [15076129](https://pubmed.ncbi.nlm.nih.gov/15076129/)]
111. Schnall R, Gordon P, Camhi E, Bakken S. Perceptions of factors influencing use of an electronic record for case management of persons living with HIV. *AIDS Care* 2011 Mar;23(3):357-365. [doi: [10.1080/09540121.2010.507745](https://doi.org/10.1080/09540121.2010.507745)] [Medline: [21347899](https://pubmed.ncbi.nlm.nih.gov/21347899/)]
112. Seifer DB, Frazier LM, Grainger DA. Disparity in assisted reproductive technologies outcomes in black women compared with white women. *Fertil Steril* 2008 Nov;90(5):1701-1710. [doi: [10.1016/j.fertnstert.2007.08.024](https://doi.org/10.1016/j.fertnstert.2007.08.024)] [Medline: [17980873](https://pubmed.ncbi.nlm.nih.gov/17980873/)]
113. Seifer DB, Zackula R, Grainger DA, Society for Assisted Reproductive Technology Writing Group Report. Trends of racial disparities in assisted reproductive technology outcomes in black women compared with white women: Society for Assisted Reproductive Technology 1999 and 2000 vs. 2004-2006. *Fertil Steril* 2010 Feb;93(2):626-635. [doi: [10.1016/j.fertnstert.2009.02.084](https://doi.org/10.1016/j.fertnstert.2009.02.084)] [Medline: [19368916](https://pubmed.ncbi.nlm.nih.gov/19368916/)]
114. Sequist TD, Cullen T, Ayanian JZ. Information technology as a tool to improve the quality of American Indian health care. *Am J Public Health* 2005 Dec;95(12):2173-2179. [doi: [10.2105/AJPH.2004.052985](https://doi.org/10.2105/AJPH.2004.052985)] [Medline: [16257947](https://pubmed.ncbi.nlm.nih.gov/16257947/)]
115. Sequist TD, Cullen T, Hays H, Taulii MM, Simon SR, Bates DW. Implementation and use of an electronic health record within the Indian Health Service. *J Am Med Inform Assoc* 2007;14(2):191-197 [FREE Full text] [doi: [10.1197/jamia.M2234](https://doi.org/10.1197/jamia.M2234)] [Medline: [17213495](https://pubmed.ncbi.nlm.nih.gov/17213495/)]
116. Serou A, Quintero R. Media influence on awareness and utilization of assisted reproduction technology in Hispanic populations. *Fertil Steril* 2011 Mar 1;95(3):e15; author reply e16. [doi: [10.1016/j.fertnstert.2010.12.030](https://doi.org/10.1016/j.fertnstert.2010.12.030)] [Medline: [21269618](https://pubmed.ncbi.nlm.nih.gov/21269618/)]
117. Shea S, IDEATel Consortium. The Informatics for Diabetes and Education Telemedicine (IDEATel) project. *Trans Am Clin Climatol Assoc* 2007 Jan;118:289-304. [Medline: [18528511](https://pubmed.ncbi.nlm.nih.gov/18528511/)]
118. Shea S, Starren J, Weinstock RS, Knudson PE, Teresi J, Holmes D, et al. Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel) Project: rationale and design. *J Am Med Inform Assoc* 2002;9(1):49-62 [FREE Full text] [Medline: [11751803](https://pubmed.ncbi.nlm.nih.gov/11751803/)]

119. Shields AE, Shin P, Leu MG, Levy DE, Betancourt RM, Hawkins D, et al. Adoption of health information technology in community health centers: results of a national survey. *Health Aff (Millwood)* 2007;26(5):1373-1383. [doi: [10.1377/hlthaff.26.5.1373](https://doi.org/10.1377/hlthaff.26.5.1373)] [Medline: [17848448](https://pubmed.ncbi.nlm.nih.gov/17848448/)]
120. Singh H, Fox SA, Petersen NJ, Shethia A, Street RL. Older patients' enthusiasm to use electronic mail to communicate with their physicians: cross-sectional survey. *J Med Internet Res* 2009;11(2):e18 [FREE Full text] [doi: [10.2196/jmir.1143](https://doi.org/10.2196/jmir.1143)] [Medline: [19632972](https://pubmed.ncbi.nlm.nih.gov/19632972/)]
121. Smith LV, Rudy ET, Javanbakht M, Uniyal A, Sy LS, Horton T, et al. Client satisfaction with rapid HIV testing: comparison between an urban sexually transmitted disease clinic and a community-based testing center. *AIDS Patient Care STDS* 2006 Oct;20(10):693-700. [doi: [10.1089/apc.2006.20.693](https://doi.org/10.1089/apc.2006.20.693)] [Medline: [17052139](https://pubmed.ncbi.nlm.nih.gov/17052139/)]
122. Sobel RM, Paasche-Orlow MK, Waite KR, Rittner SS, Wilson EA, Wolf MS. Asthma 1-2-3: a low literacy multimedia tool to educate African American adults about asthma. *J Community Health* 2009 Aug;34(4):321-327. [doi: [10.1007/s10900-009-9153-9](https://doi.org/10.1007/s10900-009-9153-9)] [Medline: [19353250](https://pubmed.ncbi.nlm.nih.gov/19353250/)]
123. Sorensen L, Gavier M, Hellesø R. Latina breast cancer survivors informational needs: information partners. *Stud Health Technol Inform* 2009;146:727. [Medline: [19592948](https://pubmed.ncbi.nlm.nih.gov/19592948/)]
124. Stanley A, DeLia D, Cantor JC. Racial disparity and technology diffusion: the case of cardioverter defibrillator implants, 1996-2001. *J Natl Med Assoc* 2007 Mar;99(3):201-207. [Medline: [17393943](https://pubmed.ncbi.nlm.nih.gov/17393943/)]
125. Stelger G, Samkoff J, Karoullas J. A program of interventions designed to increase mammography rates in women ages 50 years and older for an underserved racial minority. *J Health Hum Serv Adm* 2003 Jan;26(3):336-349. [Medline: [15704637](https://pubmed.ncbi.nlm.nih.gov/15704637/)]
126. Stern RE, Yueh B, Lewis C, Norton S, Sie KC. Recent epidemiology of pediatric cochlear implantation in the United States: disparity among children of different ethnicity and socioeconomic status. *Laryngoscope* 2005 Jan;115(1):125-131. [doi: [10.1097/01.mlg.0000150698.61624.3c](https://doi.org/10.1097/01.mlg.0000150698.61624.3c)] [Medline: [15630380](https://pubmed.ncbi.nlm.nih.gov/15630380/)]
127. Sultan-Khan LP. African American Women's Online Evaluation of the Breast Cancer Awareness and Prevention Portal of the www.divahealth.org Website: Using Personal-Level Data and Website Ratings to Tailor and Improve the Portal [dissertation]. New York, NY: Columbia University; 2010.
128. Taylor KL, Davis JL, Turner RO, Johnson L, Schwartz MD, Kerner JF, et al. Educating African American men about the prostate cancer screening dilemma: a randomized intervention. *Cancer Epidemiol Biomarkers Prev* 2006 Nov;15(11):2179-2188 [FREE Full text] [doi: [10.1158/1055-9965.EPI-05-0417](https://doi.org/10.1158/1055-9965.EPI-05-0417)] [Medline: [17119044](https://pubmed.ncbi.nlm.nih.gov/17119044/)]
129. Thompson D, Baranowski J, Cullen K, Baranowski T. Development of a theory-based internet program promoting maintenance of diet and physical activity change to 8-year-old African American girls. *Comput Educ* 2007;48(3):446-459.
130. Thornberry J, Bhaskar B, Krulewitch CJ, Wesley B, Hubbard ML, Das A, et al. Audio computerized self-report interview use in prenatal clinics: audio computer-assisted self interview with touch screen to detect alcohol consumption in pregnant women: application of a new technology to an old problem. *Comput Inform Nurs* 2002;20(2):46-52. [Medline: [11984122](https://pubmed.ncbi.nlm.nih.gov/11984122/)]
131. Thornberry JS, Murray KB, El-Khorazaty MN, Kiely M. Acceptance, Communication Mode and Use of Audio Computer-Assisted Self Interview Using Touchscreen to Identify Risk Factors among Pregnant Minority Women. *Methods Rep RTI Press* 2010 Jan 1;15:1001. [doi: [10.3768/rtipress.2010.mr.0015.1001](https://doi.org/10.3768/rtipress.2010.mr.0015.1001)] [Medline: [21637729](https://pubmed.ncbi.nlm.nih.gov/21637729/)]
132. Tirado M. Role of mobile health in the care of culturally and linguistically diverse US populations. *Perspect Health Inf Manag* 2011 Jan;8:1e. [Medline: [21307988](https://pubmed.ncbi.nlm.nih.gov/21307988/)]
133. Tomita MR, Mann WC, Fraas LF, Burns LL. Racial differences of frail elders in assistive technology. *Assist Technol* 1997 Jan;9(2):140-151. [doi: [10.1080/10400435.1997.10132305](https://doi.org/10.1080/10400435.1997.10132305)] [Medline: [10177451](https://pubmed.ncbi.nlm.nih.gov/10177451/)]
134. Tran BQ, Buckley KM, Bertera EM, Gonzales PL. Benefits & barriers to adoption of health IT in an elderly low-income, minority community-based environment. *Conf Proc IEEE Eng Med Biol Soc* 2009;2009:6441-6443. [doi: [10.1109/IEMBS.2009.5333537](https://doi.org/10.1109/IEMBS.2009.5333537)] [Medline: [19964428](https://pubmed.ncbi.nlm.nih.gov/19964428/)]
135. Trief PM, Teresi JA, Eimicke JP, Shea S, Weinstock RS. Improvement in diabetes self-efficacy and glycaemic control using telemedicine in a sample of older, ethnically diverse individuals who have diabetes: the IDEATel project. *Age Ageing* 2009 Mar;38(2):219-225 [FREE Full text] [doi: [10.1093/ageing/afn299](https://doi.org/10.1093/ageing/afn299)] [Medline: [19171951](https://pubmed.ncbi.nlm.nih.gov/19171951/)]
136. Valdez A, Banerjee K, Fernandez M, Ackerson L. Impact of a multimedia breast cancer education intervention on use of mammography by low-income Latinas. *J Cancer Educ* 2001 Jan;16(4):221-224. [doi: [10.1080/08858190109528777](https://doi.org/10.1080/08858190109528777)] [Medline: [11848671](https://pubmed.ncbi.nlm.nih.gov/11848671/)]
137. Vargas PA, Robles E, Harris J, Radford P. Using information technology to reduce asthma disparities in underserved populations: a pilot study. *J Asthma* 2010 Oct;47(8):889-894. [doi: [10.3109/02770903.2010.497887](https://doi.org/10.3109/02770903.2010.497887)] [Medline: [20846082](https://pubmed.ncbi.nlm.nih.gov/20846082/)]
138. Vincent D, Clark L, Zimmer LM, Sanchez J. Using focus groups to develop a culturally competent diabetes self-management program for Mexican Americans. *Diabetes Educ* 2006;32(1):89-97. [doi: [10.1177/0145721705284372](https://doi.org/10.1177/0145721705284372)] [Medline: [16439497](https://pubmed.ncbi.nlm.nih.gov/16439497/)]
139. Walker EA, Schechter CB, Caban A, Basch CE. Telephone intervention to promote diabetic retinopathy screening among the urban poor. *Am J Prev Med* 2008 Mar;34(3):185-191. [doi: [10.1016/j.amepre.2007.11.020](https://doi.org/10.1016/j.amepre.2007.11.020)] [Medline: [18312805](https://pubmed.ncbi.nlm.nih.gov/18312805/)]
140. Waters EA, Sullivan HW, Finney Rutten LJ. Cancer prevention information-seeking among Hispanic and non-Hispanic users of the National Cancer Institute's Cancer Information Service: trends in telephone and LiveHelp use. *J Health Commun* 2009;14(5):476-486. [doi: [10.1080/10810730903032952](https://doi.org/10.1080/10810730903032952)] [Medline: [19657927](https://pubmed.ncbi.nlm.nih.gov/19657927/)]
141. Weinstock RS, Izquierdo R, Golan R, Palmas W, Teresi JA, Eimicke JP, IDEATel Consortium. Lipid treatment in ethnically diverse underserved older adults with diabetes mellitus: statin use, goal attainment, and health disparities in the informatics

- for diabetes education and telemedicine project. *J Am Geriatr Soc* 2010 Feb;58(2):401-402. [doi: [10.1111/j.1532-5415.2009.02705.x](https://doi.org/10.1111/j.1532-5415.2009.02705.x)] [Medline: [20370876](#)]
142. Weinstock RS, Teresi JA, Golland R, Izquierdo R, Palmas W, Eimicke JP, IDEATel Consortium. Glycemic control and health disparities in older ethnically diverse underserved adults with diabetes: five-year results from the Informatics for Diabetes Education and Telemedicine (IDEATel) study. *Diabetes Care* 2011 Feb;34(2):274-279. [doi: [10.2337/dc10-1346](https://doi.org/10.2337/dc10-1346)] [Medline: [21270184](#)]
 143. West SP, Laguna C, Trief PM, Izquierdo R, Weinstock RS. Goal setting using telemedicine in rural underserved older adults with diabetes: experiences from the informatics for diabetes education and telemedicine project. *Telemed J E Health* 2010 May;16(4):405-416. [doi: [10.1089/tmj.2009.0136](https://doi.org/10.1089/tmj.2009.0136)] [Medline: [20507198](#)]
 144. Wilkin HA, Valente TW, Murphy S, Cody MJ, Huang G, Beck V. Does entertainment-education work with Latinos in the United States? Identification and the effects of a telenovela breast cancer storyline. *J Health Commun* 2007;12(5):455-469. [doi: [10.1080/10810730701438690](https://doi.org/10.1080/10810730701438690)] [Medline: [17710596](#)]
 145. Wood RY, Duffy ME, Morris SJ, Carnes JE. The effect of an educational intervention on promoting breast self-examination in older African American and Caucasian women. *Oncol Nurs Forum* 2002 Aug;29(7):1081-1090. [doi: [10.1188/02.ONF.1081-1090](https://doi.org/10.1188/02.ONF.1081-1090)] [Medline: [12183756](#)]
 146. Wood FB, Sahali R, Press N, Burroughs C, Mala TA, Siegel ER, et al. Tribal connections health information outreach: results, evaluation, and challenges. *J Med Libr Assoc* 2003 Jan;91(1):57-66. [Medline: [12568158](#)]
 147. Woolf SH, Johnson RE, Fryer GE, Rust G, Satcher D. The health impact of resolving racial disparities: an analysis of US mortality data. *Am J Public Health* 2004 Dec;94(12):2078-2081. [Medline: [15569956](#)]
 148. Wright E, Fortune T, Juzang I, Bull S. Text messaging for HIV prevention with young Black men: formative research and campaign development. *AIDS Care* 2011 May;23(5):534-541. [doi: [10.1080/09540121.2010.524190](https://doi.org/10.1080/09540121.2010.524190)] [Medline: [21287416](#)]
 149. Yost KJ, Webster K, Baker DW, Choi SW, Bode RK, Hahn EA. Bilingual health literacy assessment using the Talking Touchscreen/la Pantalla Parlanchina: Development and pilot testing. *Patient Educ Couns* 2009 Jun;75(3):295-301. [doi: [10.1016/j.pec.2009.02.020](https://doi.org/10.1016/j.pec.2009.02.020)] [Medline: [19386462](#)]
 150. Shaw SJ. The logic of identity and resemblance in culturally appropriate health care. *Health (London)* 2010 Sep;14(5):523-544. [doi: [10.1177/1363459309360973](https://doi.org/10.1177/1363459309360973)] [Medline: [20801998](#)]
 151. Crilly JF, Keefe RH, Volpe F. Use of electronic technologies to promote community and personal health for individuals unconnected to health care systems. *Am J Public Health* 2011 Jul;101(7):1163-1167. [doi: [10.2105/AJPH.2010.300003](https://doi.org/10.2105/AJPH.2010.300003)] [Medline: [21566023](#)]
 152. IQ Solutions, Inc. US Department of Health and Human Services, Office of Minority Health. 2001 Mar. National Standards for Culturally and Linguistically Appropriate Services in Health Care: Executive Summary URL: <http://minorityhealth.hhs.gov/assets/pdf/checked/executive.pdf> [accessed 2012-05-24] [WebCite Cache ID 67uH14pTo]
 153. HHS Interagency Workgroup for the NHQR/NHDR. Agency for Healthcare Research and Quality. 2011 Mar. 2010 National Healthcare Disparities Report URL: <http://www.ahrq.gov/qual/nhdr10/nhdr10.pdf>
 154. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Agency for Healthcare Research and Quality. 2008 Nov. Barriers and Drivers of Health Information Technology Use for the Elderly, Chronically Ill, and Underserved. Evidence Report/Technology Assessment No. 175. AHRQ Publication No. 09-E004 URL: <http://www.ahrq.gov/downloads/pub/evidence/pdf/hitbarriers/hitbar.pdf> [accessed 2012-05-22] [WebCite Cache ID 67r12bQBw]
 155. Gruman JC. Making health information technology sing for people with chronic conditions. *Am J Prev Med* 2011 May;40(5 Suppl 2):S238-S240. [doi: [10.1016/j.amepre.2011.02.003](https://doi.org/10.1016/j.amepre.2011.02.003)] [Medline: [21521600](#)]
 156. Rudd RE, Anderson E, Oppenheimer S, Nath C. Health literacy: an update of medical and public health literature. In: Comings JP, Garner BP, Smith CA, editors. *Comings J, Garner B, Smith C, editors. Review of Adult Learning and Literacy, volume 7: Connecting Research, Policy, and Practice (National Center for the Study of Adult Learning and Literacy Series)*. New York, NY: Lawrence Erlbaum; 2007:175-204.
 157. Agency for Healthcare Research and Quality. 2011 Mar. Health Literacy Interventions and Outcomes: An Updated Systematic Review URL: <http://www.ahrq.gov/clinic/tp/lituptp.htm> [accessed 2012-05-24] [WebCite Cache ID 67uH2krGY]
 158. Barrera M, Castro FG, Strycker LA, Toobert DJ. Cultural adaptations of behavioral health interventions: a progress report. *J Consult Clin Psychol* 2012 Jan 30. [doi: [10.1037/a0027085](https://doi.org/10.1037/a0027085)] [Medline: [22289132](#)]
 159. Brown M. LGBT aging and rhetorical silence. *Sex Res Social Policy* 2009;6(4):65-78.
 160. Daley A, Solomon S, Newman PA, Mishna F. Traversing the margins: intersectionalities in the bullying of lesbian, gay, bisexual and transgender youth. *J Gay Lesbian Soc Serv* 2008;19(3-4):9-29. [doi: [10.1080/10538720802161474](https://doi.org/10.1080/10538720802161474)]
 161. Ivanitskaya L, O'Boyle I, Casey AM. Health information literacy and competencies of information age students: results from the interactive online Research Readiness Self-Assessment (RRSA). *J Med Internet Res* 2006 Jan;8(2):e6 [FREE Full text] [doi: [10.2196/jmir.8.2.e6](https://doi.org/10.2196/jmir.8.2.e6)] [Medline: [16867969](#)]
 162. Chavez EL, Oetting ER. A critical incident model for considering issues in cross-cultural research. Failures in cultural sensitivity. *Int J Addict* 1995 May;30(7):863-874. [Medline: [7558474](#)]
 163. Zanchetta MS, Poureslami IM. Health literacy within the reality of immigrants' culture and language. *Can J Public Health* 2006;97 Suppl 2:S26-S30. [Medline: [16805158](#)]

164. DeCarlo P. AIDS Partnership California. 2003 Oct. Completing the Circle: Designing HIV Prevention Programs for Persons of Color With HIV URL: <http://www.aidspartnershipca.org/assets/CompletingTheCircleBW.pdf> [accessed 2012-05-24] [WebCite Cache ID 67uH8evwR]
165. Chang VW, Lauderdale DS. Fundamental cause theory, technological innovation, and health disparities: the case of cholesterol in the era of statins. *J Health Soc Behav* 2009 Sep;50(3):245-260. [Medline: [19711804](#)]
166. Mangold WG, Faulds DJ. Social media: the new hybrid element of the promotion mix. *Bus Horiz* 2009 Jul;52(4):357-365.
167. Pearson ML, Mattke S, Shaw R, Ridgely MS, Wiseman SH. Agency for Healthcare Research and Quality. 2007 Nov. Patient Self-Management Support Programs: An Evaluation URL: <http://www.ahrq.gov/qual/ptmgmt/> [accessed 2012-05-24] [WebCite Cache ID 67uH7LUiS]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality

AIDS: acquired immunodeficiency syndrome

EHR: electronic health record

HIV: human immunodeficiency virus

IT: information technology

Edited by G Eysenbach; submitted 26.02.12; peer-reviewed by L Kvasny; comments to author 19.03.12; revised version received 28.03.12; accepted 25.04.12; published 31.05.12.

Please cite as:

Montague E, Perchonok J

Health and Wellness Technology Use by Historically Underserved Health Consumers: Systematic Review

J Med Internet Res 2012;14(3):e78

URL: <http://www.jmir.org/2012/3/e78/>

doi: [10.2196/jmir.2095](https://doi.org/10.2196/jmir.2095)

PMID: [22652979](https://pubmed.ncbi.nlm.nih.gov/22652979/)

©Enid Montague, Jennifer Perchonok. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 31.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Interaction Patterns of Nurturant Support Exchanged in Online Health Social Networking

Katherine Y Chuang¹, BS, MLIS; Christopher C Yang¹, PhD

College of Information Science and Technology, Drexel University, Philadelphia, PA, United States

Corresponding Author:

Katherine Y Chuang, BS, MLIS

College of Information Science and Technology

Drexel University

3141 Chestnut St

Philadelphia, PA, 19104

United States

Phone: 1 215 895 2474

Fax: 1 215 895 2494

Email: katychuang@drexel.edu

Abstract

Background: Expressing emotion in online support communities is an important aspect of enabling e-patients to connect with each other and expand their social resources. Indirectly it increases the amount of support for coping with health issues. Exploring the supportive interaction patterns in online health social networking would help us better understand how technology features impacts user behavior in this context.

Objective: To build on previous research that identified different types of social support in online support communities by delving into patterns of supportive behavior across multiple computer-mediated communication formats. Each format combines different architectural elements, affecting the resulting social spaces. Our research question compared communication across different formats of text-based computer-mediated communication provided on the MedHelp.org health social networking environment.

Methods: We identified messages with nurturant support (emotional, esteem, and network) across three different computer-mediated communication formats (forums, journals, and notes) of an online support community for alcoholism using content analysis. Our sample consisted of 493 forum messages, 423 journal messages, and 1180 notes.

Results: Nurturant support types occurred frequently among messages offering support (forum comments: 276/412 messages, 67.0%; journal posts: 65/88 messages, 74%; journal comments: 275/335 messages, 82.1%; and notes: 1002/1180 messages, 84.92%), but less often among messages requesting support. Of all the nurturing supports, emotional (ie, encouragement) appeared most frequently, with network and esteem support appearing in patterns of varying combinations. Members of the Alcoholism Community appeared to adapt some traditional face-to-face forms of support to their needs in becoming sober, such as provision of encouragement, understanding, and empathy to one another.

Conclusions: The computer-mediated communication format may have the greatest influence on the supportive interactions because of characteristics such as audience reach and access. Other factors include perception of community versus personal space or purpose of communication. These results lead to a need for further research.

(*J Med Internet Res* 2012;14(3):e54) doi:[10.2196/jmir.1824](https://doi.org/10.2196/jmir.1824)

KEYWORDS

Social support; social media; alcoholism

Introduction

The Internet is a tool that can quickly connect people to each other, forming niche communities that house conversations. Many people go online for communicative or social reasons,

as well as tailored information, as they face important decisions [1-4]. In 2009, the Pew Internet & American Life Project found that e-patients wanted to access user-generated or “just-in-time someone-like-me” health information such as newsgroups, blogs, social networking sites, or updates [4]. It is very likely that they are looking for compassion or experiential knowledge.

E-patients are Internet users who seek, share, and sometimes create information about health and wellness [4]. They benefit from sharing their experiences, discussing medical information, and exchanging social support. Social support through peer communication can enable e-patients to cope with stress, and it also increase access to information [5,6]. Emotional support, which is one type of social support, can lead to improved health outcomes such as easing adjustment to cancer [7]. Many users join online support groups for a sense of community around those who have similar experiences [7,8]. In fact, some researchers suggest that online communities become surrogate families of e-patients, where members share common problems and help each other toward mutual goals through good times and bad [7,8].

Online communities exist across websites built upon social technologies, such as bulletin boards and mailing lists, and social network sites. Social network sites are different from previous text-based communication formats in their emphasis on the website representing relationships between users and architectural elements that encourage interpersonal relationships [9]. Social network sites enable users to find each other and build connections using profile pages and a spectrum of private and public communication tools [10,11]. These social technologies are particularly useful for health services in enabling the creation of Web interventions to help heavy drinkers [12-14].

Studying the patterns of nurturing interactions within a support community gives useful insight into users' information and emotional needs from an online support group. Each computer-mediated communication format has a unique combination of characteristics that affect the interaction patterns of nurturant support. Understanding the relation between the technology and how it affects user behavior could help both end users and designers of these online communities. Supposing that one computer-mediated communication format may have a higher association with a particular type of question, this could be useful in the design of Web interventions that incorporate online communities [15]. Studies such as the present one that report social interaction patterns in an online community could reveal areas of improvement for software that serves as the backbone of online communities. The objective of our research was to identify the types of nurturing social support in an online health care social networking site across three different text-based communication tools: discussion forums, personal journals, and notes. This study presents empirical data as evidence that patterns of nurturing support vary among the different online tools. We discuss our results in conjunction with theories to explain possible causes of support behavior patterns.

Related Work

Previous research documented examples where e-patients needed online social support beyond technical support and information from health providers (eg, doctor visit reminders); however, these studies were limited on the technical front due to overlooking specific characteristics in the computer-mediated communication format as a factor of support group behavior. Computer-mediated communication enables interpersonal

communication in a public environment, but end users often conduct their conversations on this platform as a private space [9,16-18]. This dilemma presents the motivation to compare supportive communication across multiple formats.

Online Social Support

Support communities are likely to exist in a health domain when patients experience stress or confusion at a personal level [19]. E-patients' information needs often coexist with other conditions (eg, receiving unfortunate news or requiring a behavior change). For example, a patient with liver cancer may have learned that the cancer was a result of overdrinking over many years. When shifting to a sober lifestyle, the patient may find general social support from friends to be helpful in preventing a relapse [20].

Of the different social support types available (informational, nurturant, and instrumental), nurturant support is known as a more intimate type of support, built upon trust and more frequent interactions [21,22]. It includes expressions of caring for someone without necessarily seeing tangible efforts. Social network sites are online spaces where users can work on creating intimate relationships with each other using the friending and messaging features, thus exchanging social support [10,23]. Expressing emotion is an important component in the daily management of relationships. Unlike in other online communities, users of social network sites tend to expect to gratify their social-emotional needs rather than informational needs [22]—for instance, positive comments between friends on general social network sites such as MySpace [23]. Close friends are sources of emotional support that can help with coping in difficult times or to improve mental health [24-26]. People with strong ties often communicate through many channels [24,27], but also with more emotional content than when communicating with strangers. We suspect that the different communication formats can reveal a spectrum of nurturant support types.

Positive outcomes from participation in self-help groups include (1) sharing information such as ideas, facts, and resources, (2) engaging in dialogue to reveal multiple perspectives, (3) discussing taboo subjects, (4) being “all in the same boat” with others, (5) experiencing mutual support, (6) engaging in problem solving and rehearsing, (7) overcoming alienation and isolation, (8) engaging in catharsis, (9) taking on the role of helper, (10) developing inspiration and hope, (11) developing social networks, and (12) assisting more people less expensively [8,19,28,29]. Support community members often share personal experiences, which may also include personal information. Emotional support is a valuable element of social support and in helping support seekers in coping with health problems [19,28,29]. Peer communication such as establishing social norms or finding role models and sharing feelings can also play a role in facilitating new health habits, such as quitting smoking [30]. Because of the nature and helpfulness of emotional support, we explored the nurturing types (esteem, network, and emotional) across the various computer-mediated communication formats.

Medical Environment of Sharing Nurturant Support

While similar studies over the past decade also identified the content within online support communities across several health conditions, they did not focus on the impact of technology on the communication. Table 1 [6,19,21,28, 31-36] shows a brief

list of these related studies. For instance, they compared types of social support exchanged across email lists and discussion groups [19,21,28,31] and bulletin boards [6,32-36]. These earlier forms of social technologies lacked some features of current social media computer-mediated communication formats, particularly profile pages.

Table 1. Previous studies of online support in health-related social networking sites.

Study	Data set
McCormack [6]	Eating disorder
Preece [19]	Torn knee ligament, 500 messages, April 1996–April 1997
Bambina [21]	Support OnLine Cancer Forum, 84 members, 1149 messages
Braithwaite et al [28]	Support Network for disabilities, 42 users, 1472 messages
Meier et al [31]	Association of Cancer Online Resources
Cunningham et al [32]	Alcoholism, 10 months, 474 posts (moderated)
Coursaris et al [33]	HIV/AIDS ^a , 5000 messages
Eichhorn [34]	5 eating disorder message boards, 490 messages
Pfeil and Zaphiris [35]	Depression/seniors, discussion forum
Selby et al [36]	StopSmokingCenter.net (moderated) (November 6, 2004–May 15, 2007)

^a Human immunodeficiency virus/acquired immunodeficiency syndrome.

Previous studies of online support communities also identified various types of support that they provided [6,19,21,28,32-35], and compared online support with face-to-face empathy [35] and with other types of communities [19]. They gave examples of benefits from participation [8,37]; however, these studies did not discuss their findings in light of communication format characteristics or looking at the relationship between the use of the tool and the tool itself. We were interested in a more in-depth study of social software features that affect social interaction patterns.

Some of these studies focused on evaluating support communities [32,36], while many of the other studies aimed to describe the content of the support communities [19,21,28,35]. Some studies also investigated the types of social support exchanged using Cutrona and Suhr's [38] social support framework [33,34]. Evaluation could be useful in understanding what users need from the online support community.

Computer-Mediated Communication Through Social Network Sites

We viewed social support exchanges in an online community from an architectural perspective, where the site design affects user behavior [9,39,40]. Certain features, such as privacy level settings, suggest acceptable behaviors to promote “the development of particular culture or behaviors and identity presentation;” however, users will customize them to improve their social interactions [9,39]. This might be a cause for concern because electronic media lack clear boundaries between traditionally public and private spaces [9,40]. In the physical world, the walls of offices and houses clearly separate distinct situations, and gates section off personal property. However, the Internet blurs the separation between public and private information in the online space [9]. In social network sites, users must balance their private and public selves, especially when

conducting social interactions with each other. For example, Facebook is similar to a greenhouse, with its publicly open structure and many communication tools that members use to leave social cues for each other [9]. Public comments and other communication can signal the “strength and context of a relationship” [39].

Social network sites encourages disparate individuals to connect, communicate, and take action, which fosters interaction that is primarily interpersonal [9,41]. Social network sites provide the capability for users to represent themselves with an online presence (identity information) that contains shareable personal information, such as their birthday, preferences, photographs, and writings, and can assist in developing common ground and facilitate initial interactions [11,41]. Convenient features allow users to form and maintain online network “friends,” where, if one user invites another user with a friend request and is accepted, a relationship is established on the website [10]. Friends can communicate through social network sites in several ways, including private and public messaging systems [9,10,23,26,41]. Studying these interactions (eg, the length, frequency, and content of these comments) rather than explicit articulations on profile pages can reveal the conversational profile of each relationship [9,39].

Social network site interactions are founded upon norms of everyday face-to-face interaction, but when they are adapted to the online setting, the distinction between public and private spheres is blurred [17]. The space experience comes from relations with others [42,43]. In neutral spaces such as urban spaces that offer public gatherings, a group's main activity is informal conversation. Public spaces such as parks are communal and have certain purposes of use compared with a private space such as a home. Communication tools are designed with a spectrum of privacy options, where each

computer-mediated communication format may project toward users a different perspective of intimacy, and therefore lead to different social supports being offered and sought. For instance, users might post questions with factual answers in a forum that all other users can easily notice and access, for a better chance to receive a quick answer from a broad audience. On the other hand, users who seek encouragement may prefer using a personal journal to limit support requests to a smaller group of closely related users.

The research literature also suggests that, regardless of context, software features affect user behavior in online communities [9,16]. More specifically, elements within each computer-mediated communication format affect the resulting social interactions. In this study, we observed interactions in the forum, across user journals, and in profile posts on MedHelp. The forum is similar to a community hall in the sense that the space belongs to the community. In that community space, anyone can post in, comment on, or read discussion threads. All forum post and comments of a medical support community are available in the corresponding support community page. The threads are sorted by the dates when they are posted. The access authorization is managed by the community administration. In contrast to the forum, journals and notes are

similar to friends sitting in a coffee shop and having a personal conversation. All journal messages and notes are posted on a user's personal profile page instead of on a community page. There is no single aggregate page listing multiple journal authors or multiple note recipients. Instead, users must go through the personal profile page of an individual user to post a message. The user (or owner) of the personal profile page can control access to journals and notes to "Everyone," "Only my friends," or "Only me." In this personal space, any passerby would be near enough to overhear comments but would immediately interrupt the conversation. The distinction between the sense of communal space and personal space derives from certain architectural elements, which determine the openness to new messages (see Table 2). In a way, these computer-mediated communication formats are similar to MySpace's and Facebook's privacy features that allow a user to put a fence around personal property. A unique feature in these health social network sites allows users to add optional health data such as daily weight to forum and journal posts to share with others. Reading new messages also varies across each format; new forum messages are listed collectively on the support community page, and journal and notes are listed on each user's profile page.

Table 2. Comparison of architectural elements in computer-mediated communication formats in MedHelp.

Format	Architectural element				
	Authoring new posts	Commenting	Access and notifications	Privacy settings	Other features
Forum	Any user	Any user	Public Forum Page	Public	Add Tags, Select topic, Add to watch list, Show Ticker
Journal	Owner of journal	Friends (depending on settings)	Profile Page	Everyone, Only my friends, Only Me	Add Tags, Show Ticker, Add Photo
Notes	Friends	None	Profile Page	Everyone, Only my friends, Only Me	Add as friend

By exploring the range of nurturing behaviors displayed across multiple computer-mediated communication formats of an online health supportive community, we can understand how the software features affect communication between individuals of the community. Previous research did not address the communication platform characteristics as an influence on resulting user behaviors. We used content analysis to identify themes in user-created content. We speculated that patterns of nurturant support exchanged across different social media communication formats may vary depending on the architectural elements characterizing the communication tools. In this study we tried to answer the following question: What are the differences in nurturant support types (esteem, network, and emotional) across different communication formats (forum, journal, and notes)? Results from this study would be useful for improving the design of technologies supporting online communities, because the increasing socialization of online health information will open up opportunities for future online health services [4]. By answering these research questions, we hope to better understand the link between site design and group interactions in an online health community.

Methods

We collected data from a health social media website called MedHelp [44], which has features similar to those in Facebook to help people connect with each other. It is one of the oldest patient communities on the Internet, founded in 1994 by a software developer and pharmaceutical/biotech professional, both of whom were touched personally by their families' health conditions. Their website has over 12 million visits each month. Alcoholism is one of the leading health problems in the United States, with many researchers publishing studies of the disorder. The Alcoholism Community is one of the biggest communities on MedHelp, with many user activities, which provides enough data for analysis.

Our approach consisted of data collection with a Web crawler and manual content analysis. Although the typology of social support used in this study was originally developed for short conversations between spouses rather than a support group, it has been shown to be sufficiently generalizable for use in support groups [28,33,34,38]. All five of the supracategories from Cutrona and Suhr's [38] typology were coded in our data. Of these five, we did not find instrumental support, such as the

tangible assistance of offering the recipient a loan and offering to perform an indirect task.

Study Setting: MedHelp Support Communities

MedHelp is a health-oriented social networking service where individuals can connect with other people (patients, caregivers, doctors, etc) and information resources. Its platform provides an environment for registered users to join peer support communities as members and to communicate using several tools, such as discussion forums, journals, and notes. The forum for each peer support community is unmoderated, in the sense that any MedHelp user is able to post questions to the community or respond with comments without undergoing an approval process. The journals format allows users to record

thoughts and feelings. Notes are a way for users to keep in touch with each other through their profile pages on MedHelp.

The support community page lists the most recently updated forum threads in the main section and recent activity on the sidebar, which also lists recent updates or journal comments (Figure 1). There is also a sidebar box that lists community members and links to their profile pages. Each personal profile page displays sections of the user's activity across communication tools (Figure 2). Privacy settings on an individual user's profile page can affect who can read updates and write notes. If the settings for journals and notes are set to Only my friends, then only users who are friended may view these contents. If the setting is set to Only me, only the user can see his or her own content when logged in. The content in each of these tools is organized chronologically.

Figure 1. Community activity (messages and updates) are displayed on public forum page.

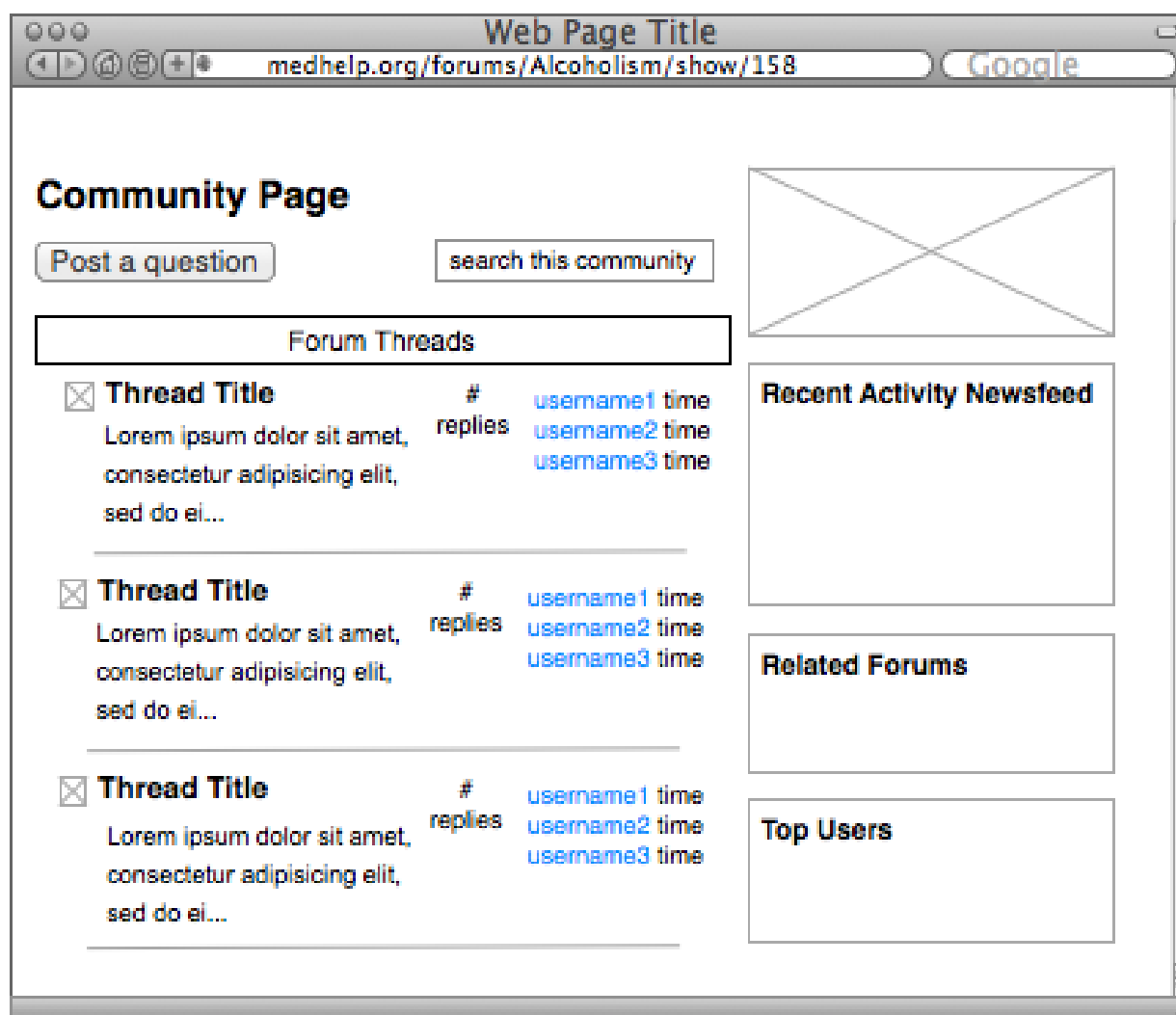
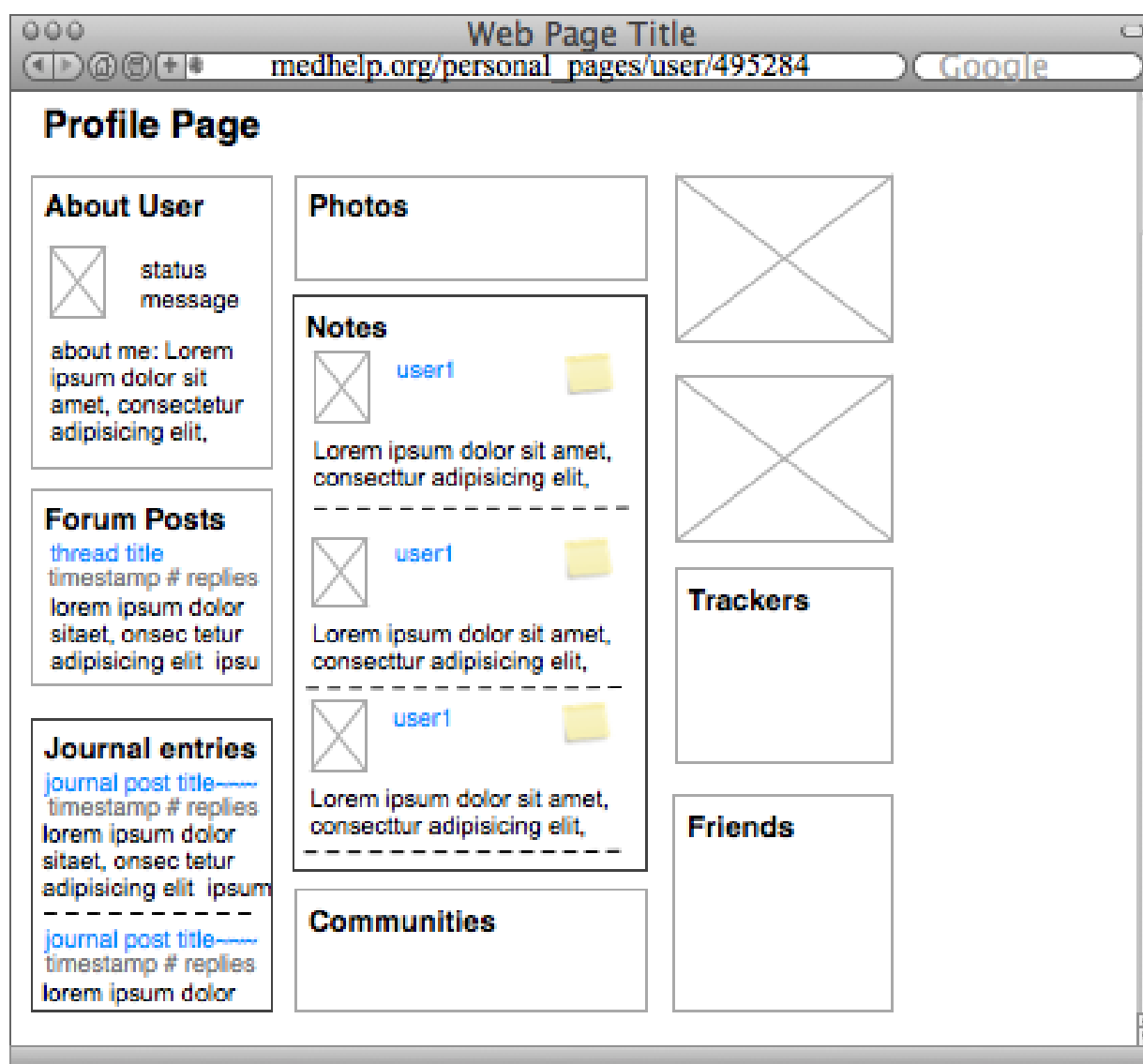


Figure 2. Friends and messages displayed on user's profile page.

How Users Message Each Other in MedHelp

The three communication tools we investigated in this study (forum, journals, and notes) are available for any MedHelp support community member to post content. Each tool varies in features for posting content (Figure 3, Figure 4, Figure 5). Users can post questions or polls to the forum, record thoughts in journals, and contact other users with notes.

To post a question in the forum, users are required to type in a title, select a topic, describe their question, and optionally add tags (Figure 3). When posting to the forum it is fairly clear that the intention is for the user to communicate with the entire support community.

A posting to a journal can include title, entry, tags, and photos, with selected privacy options (Figure 4), and is listed only on a user's profile page. The intention of posting to the journal for

MedHelp users is to record their thoughts and feelings. Only the owner of the personal profile page can initiate a thread by making a journal post, but other authorized users are able to comment on a journal post.

Notes on a user's personal profile page include type of note and the content in the note (Figure 5). Any member who is allowed to read the notes entries is allowed to post. Users can leave notes on other members' pages if they are allowed to read it. If the user is not a friend, there is an option to befriend the user. However, the owner of the personal profile page is not allowed to post to his or her own notes section.

The journal and notes features are similar to Facebook and MySpace wall posts, which allow users to communicate one-on-one. However, MedHelp is more specific on who can initiate, comment on, and post in journal and notes sections. The screen shots are shown in Figures 3–5.

Figure 3. Posting a question on the forum.

New Question - Alcoholism (Medical Community)

☒ Ask a question ☐ Post a poll * Required

Title for your question *


Topic *

Friend and Family Support ▾

Your Question *

Tags

Comma separated (for example "pregnancy, surgery, cold sore")

▶ Add Photo 

▶ Show Ticker

☒ Add to watch list

☒ Join this community

Post Question

NOTE: Any information you submit will appear to the public in our forum, so DO NOT enter your LAST name, E-MAIL ADDRESS, or any other confidential information.

[Back to Forum](#)

Figure 4. Posting in the journal.

Figure 5. Posting in notes.

Data Collection and Preparation

We developed a Web crawler in the Java language to scrape data from the MedHelp Alcoholism Community [44] and saved the data to text files. The text files were then converted to Excel spreadsheets (Microsoft Corporation, Redmond, WA, USA) for identifying themes during the descriptive content analysis.

The data we were interested in collecting included messages created by members in the Alcoholism Community. We collected publicly available messages from the discussion forums, user journals, and notes on users' profile pages. Because the forum was slightly different from journals and notes, we took different steps for each sample. To collect forum messages,

we had the crawler scrape each thread of messages and divided it into posts and comments. To collect user journal messages and profile posts, we first identified the list of users who publicly listed their memberships in the MedHelp Alcoholism Community. Then on each user's profile page, we specified that the crawler go to the link for journals to scrape publicly available entries and their corresponding comments. We also specified that the crawler scrape the profile posts. For each message the crawler identified, we attached information, if it was available, about the author, recipient, timestamp, and titles.

Content Analysis

We used descriptive content analysis to find patterns of social support in messages in the MedHelp Alcoholism Community

interactions across 3 months of data (June 19 to September 19, 2009). More than one theme could appear per message. No identifiable information, such as each participant's drinking problems, was part of the final analysis.

Our unit of analysis was at the message level. We developed definitions of social support types by reviewing examples from the related literature and matched them with themes within the data [21,38,45,46]. A pretest was conducted to determine whether the categories captured a majority of the message postings. Results of the pretest informed our decision to organize our definitions into three categories (information, nurturant, and instrumental). The interrater reliability of the coding system was determined using two coders who were both graduate students in information science with health informatics experience. They received the coding instructions and conducted the coding independently. The results were recorded onto an Excel spreadsheet. We used Cohen kappa to determine the interrater agreement ($\kappa = .719$). This indicated that a high agreement was achieved.

We found only two types of support in the data: informational support and nurturant support. The third type, instrumental support, is typically found in face-to-face support interactions and was not found in our data.

Support Types: Informational Versus Nurturant Support

Social support is the provision of psychological and tangible resources intended to benefit an individual's ability to cope with stress, such as information or statements leading the person to believe he or she is cared for. We discuss the types of social support messages used in computer-mediated support groups, organized within Cutrona and Suhr's [38] guideline for categories of informational support and nurturant support. There are two main types of support: (1) *action facilitation*, which is

intended to help stressed individuals to solve or eliminate problems causing distress, and (2) *nurturant support*, which caters to the emotional side by comforting support seekers [36,38]. Action facilitation support includes both informational support and instrumental support. Informational support could be facts, advice, information referral, personal stories, or opinion. Instrumental supports are tangible services, either direct or indirect, for improving the situation (eg, driving a friend to the hospital). In this study we focused on the second type of support occurring in different computer-mediated communication formats of an online health support community because we wanted to better understand how a specific support relates to computer-mediated communication formats and because we found more messages with nurturant support. We looked for patterns of nurturing interactions across each format to identify differences in behavior across the group. Nurturant support includes esteem support, network support, and emotional support, which are defined in more detail in the next section.

Coding Scheme: Informational Support Types

Informational support in messages conveys instructions, including (1) advice or teaching, (2) referrals to other sources of information, (3) situation appraisal, (4) stories of personal experience, and (5) opinions. Messages coded as informational support often appeared as an attempt to reduce uncertainty for the message recipient [28,34,38]. We identified these informational support types in the forums, journals, and notes.

Coding Scheme: Nurturant Support Types

Nurturant support posts provide expressions of caring or concern [30,33,34,38,47,48]. These are summarized in Table 3. Nurturant support is a more compassionate type of support whose purpose is to help the recipient with coping or relieving stress. It has been studied in a variety of online patient support communities, listed in Table 4 [6,19,21,28,31-34,37,49].

Table 3. Definitions of nurturant support types for content analysis measures.

Support type	Definition	Example
Esteem	Gives positive comments validating a recipient's self-concept, alleviating feelings of guilt as a person; includes compliments, validation, and relief from blame [36]	<i>Congratulations on your sobriety!</i>
Compliment	Conveys positive assessment toward someone or emphasizes the recipient's skills and abilities	<i>Thanks x. This is a great journal entry. Thanks for the laughs. I really needed it. My favorite one is #16.</i> <i>Good Job! I'm sooooo proud of you =0) :o) =)</i>
Validation	Recognizes need by expressing agreement with the recipient's perspective on the situation	<i>X, no need to apologize its a great post especially when it comes from your heart <3</i>
Relief of blame	Tries to alleviate the recipient's feelings of guilt about the situation	<i>Its not your fault</i> <i>Don't blame yourself</i>
Network	Focuses on messages to broaden support seekers' social network so they don't feel alone, by connecting them to others with similar situations; includes access, presence, and companions [36]	
Access	Invites new members to join conversations or offered to connect members with others having similar interests	<i>Well, I guess I wasn't much help, but I appreciate the input, and it's good to know you're not alone. Thank you x. Maybe we can help each other.</i>
Presence	Offers to spend time with the person, to be there in time of need	<i>...well my dear...please stay in touch with us here...we do care what happens to ul!:))</i>
Companions	Reminds the person of availability of companions, of others who are similar in interests or experience	<i>Just reach out and I will be there ok?</i>
Emotional	Gives expressions that support recipient's feelings or reciprocates emotion; the emphasis of this category is on supporting emotional states rather than the recipient's identity or self-concept [36,46]	<i>You're going through a rough time</i> <i>Hang in there hon.</i>
Relationship	Stresses the importance of closeness and love in relationships the recipient has with others	<i>I have missed this forum SO much! Finally back from crazy land and look forward to seeing all those familiar names comment on the questions.</i>
Physical affection	Offers physical contact, including hugs, kisses, hand-holding, shoulder patting; obviously, physical affection could not be given online, but it was often offered and conveyed verbally	<i>You deserve a big bear hug!</i>
Confidentiality	Promises to keep the recipient's problem in confidence; confidentiality is typically symbolic [28]	
Sympathy	Expresses sorrow or regret	<i>Sorry it had to happen to you.</i>
Listening	Provides attentive comments as the recipient speaks	
Understanding and empathy	Expresses understanding of the situation or discloses a personal situation that communicates similarity of one person's experiences with another's	
Encouragement	Provides messages of hope or confidence	<i>Hang in there!</i>
Prayers	In spiritual words, mentions praying, spiritual healing, or God	<i>I will keep you in my prayers.</i>

Table 4. Comparison of nurturant support types in previous studies (by reference number).

Support type	Studies and settings						
	Disability	Eating disorder	HIV/AIDS ^a	Addiction	Torn knee ligament	Depression	Cancer
Esteem	[28]	[6,33]	[32]	[31]			[49]
Compliment	[28]		[32]	[37]			
Validation	[28]		[32]			[34]	[21] ^b
Relief of blame	[28]		[32]				
Network	[28]	[6,33]	[32]	[31]		[34]	
Access	[28]		[32]				
Presence	[28]		[32]				
Companions	[28]		[32]				[21]
Emotional	[28]	[6,33]	[32]			[34]	[49]
Relationship	[28]		[32]				
Physical affection	[28]		[32]				
Confidentiality	[28]		[32]				
Sympathy	[28]		[32]				[21]
Listening	[28]		[32]				
Understanding and empathy	[28]		[32]	[31,37]	[19]		[21,49]
Encouragement	[28]		[32]		[19]	[34]	[21,49]
Prayers	[28]	[6]	[32]			[34]	[49]

^a Human immunodeficiency virus/acquired immunodeficiency syndrome.

^b This study used the label emotional support for validation support.

Results

The three samples of data consist of user-created messages from the discussion forums (n = 493), user journals (n = 423), and

notes (n = 1180). Based on the displayed structure, forum and journal messages were grouped into *posts* (ie, messages that start the thread) and *comments* to the post. The data sets encompassed 81 forum posts; 412 forum comments, 88 journal posts, and 335 journal comments (Table 5).

Table 5. Summary of samples and their sizes.

Type of message	Sample		
	Forums	Journals	Notes
Messages	493	423	1180
Posts	81	88	NA ^a
Comments	412	335	NA

^a Not applicable.

There was a range of message characteristics. A message contained on average 2.57 codes with a maximum of 10 codes per message, except for the first post of each thread, which had a maximum of 6 codes. Some messages only offered support (eg, “Have you tried Naltrexone? It is sup[p]osed to help with the cravings there are other meds that can help with it too. If all else fails, make a picture of tea and pop some popcorn and hang out with him with your ‘drink’.”), or only requested support (eg, “Hi, is there a medicine to take to stop the craving for alcoholic drink?”). We first identified informational and nurturant support in the samples (both provided and requested)

and present these findings first before presenting specific nurturant support types.

Informational Versus Nurturant Support

The contrast between informational and nurturant support was apparent among the computer-mediated communication formats. Discussion forums were more informational than the other two (journals and notes). More nurturant support was offered in notes and journal comments, whereas more informational support was offered in the forum. However, when requesting social support, users were not as likely to seek nurturant support.

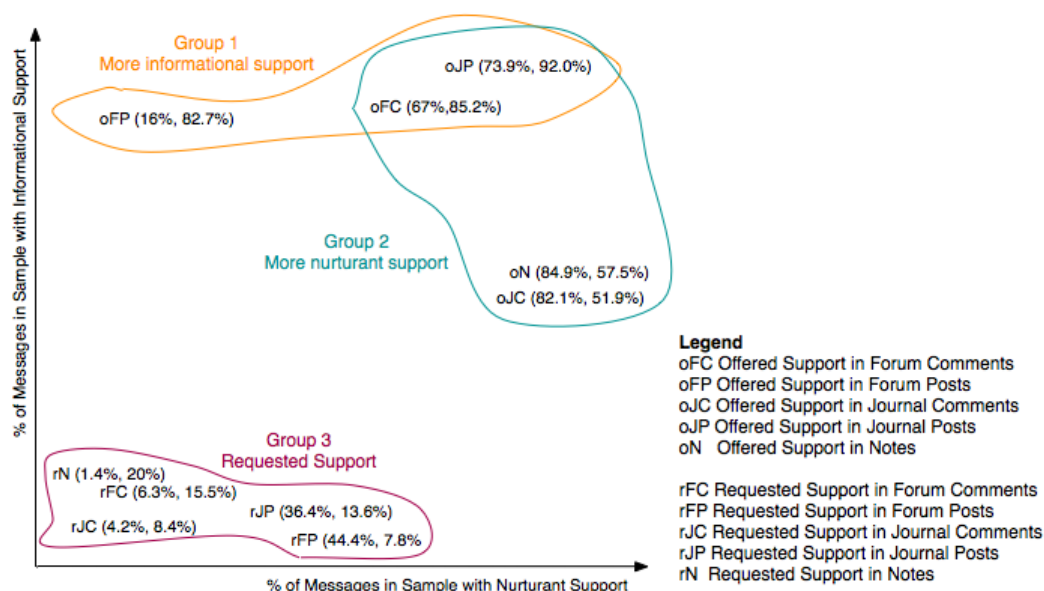
The only exception is in journal posts, where the user was more likely to seek nurturant support.

To illustrate the differences in the computer-mediated communication formats, we plotted data onto a graph (Figure 6). The points are noted in parenthesis (x,y), where x indicates the number of messages in the sample with nurturant support and y indicates the number of messages in the sample with informational support. The sample name has a prefix that signifies offered support (o) or requested support (r). Three main clusters showed up. *Group 1* contained formats where a majority of messages offered informational support (eg, advice). This type of support was found in 83% (67 of 81 messages) of forum posts, 85.2% (351 of 412 messages) of forum comments, and 92% (81 of 88 messages) of journal posts. *Group 2* contained formats where a majority of messages offered nurturant support (eg, encouragement). This type of support was found in 74% (65 of 88 messages) of journal posts, 67.0% (276 of 412 messages) of forum comments, 84.9% (1002 of 1180 messages) of notes, and 82.1% (275 of 335 messages) of

journal comments. *Group 3* contained all the formats' levels of requested information and nurturant supports, which were relatively low compared with offered supports.

Our initial analysis showed more messages with nurturant than informational support. This is an interesting type of support to investigate because it tends to be more interactive in a bidirectional way. We were interested in studying this to have a better understanding of how a specific social support can be facilitated. In our next step, we analyzed the specific types of nurturant support more carefully for each sample, and then compared results across all samples. For example, there was a link between requested support in journal posts (37%, or 32 of 88 messages with nurturant) and offered support in journal comments (82.0%, or 338 of 412 messages with nurturant). There was an interesting connection in reverse for notes, where users requested information but provided more emotional support. While forums may be seen as a question-and-answer portal for exchanging information, the portion of nurturant support was higher than expected in the comments.

Figure 6. Informational and nurturant support levels in user-created messages.



Nurturant Support Offered

Table 6 summarizes the number of messages showing nurturant support offered for each sample. After tabulating these numbers, we converted them into percentages. Figure 7 displays the

percentages and shows two types of patterns. Two nurturant support patterns emerged in messages that offered support: emotional > network > esteem (forum posts, notes) and emotional > esteem > network (forum comments, journal posts, journal comments). We explain these two patterns in the section.

Table 6. Number of messages offering nurturant support in each format.

Support type	Forums		Journals		Notes
	Posts	Comments	Posts	Comments	
Esteem	1	53	13	124	220
Network	5	18	2	17	488
Emotional	12	259	61	241	752
Total	18	330	76	382	1460

Emotional was the most commonly appearing subtype among offered nurturant support (Figure 7). Network and esteem

occurred less in comparison. In two sets (journal comments, forum comments), esteem is greater than network (esteem >

network). This pattern may reflect the compassionate nature of users who recognized the perspective of the first author, compliments, or relieving blame. In addition, journal posts also displayed more esteem than network, which may indicate their authors' awareness of their audience. Conversely, in the samples that had more network than esteem (notes, forum posts), the strategy might have been to increase communication with an emphasis on presence, access, or companionship.

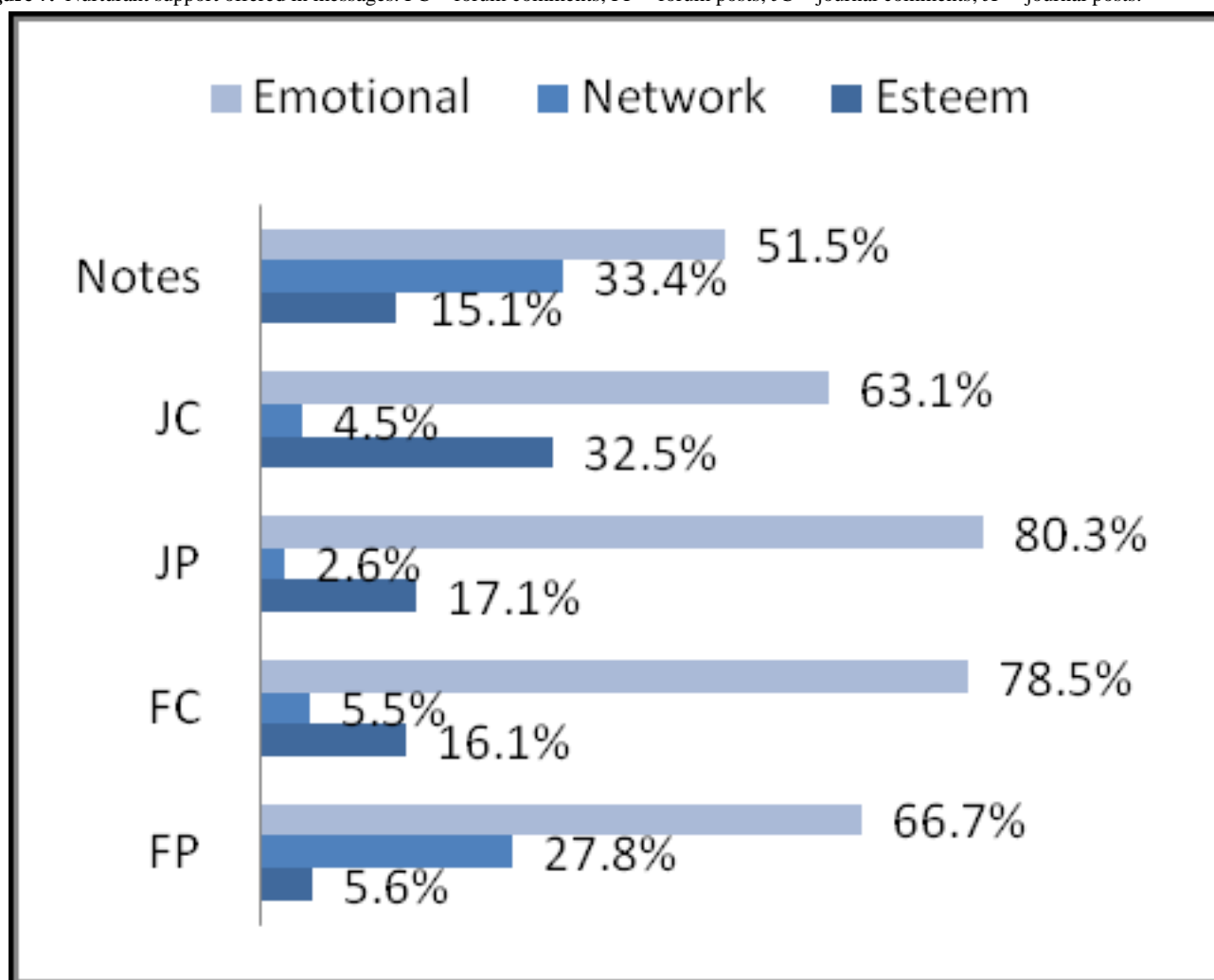
The first pattern where network is greater than esteem support might be an indication of the user's informational or emotional state while starting a thread in the forum or creating a note for a friend. The emphasis on network over esteem suggests promoting presence and involvement with the community, which could be a strategic expression for being a worthy recipient of social support. The lower amount of esteem support might suggest social status as less important than membership. The similarity between forum posts and notes may suggest that users post without expectation of a direct response.

The second pattern, where the communication formats had the pattern of more esteem than network support (emotional >

esteem > network; forum comments, journal posts, journal comments), may suggest that these formats are more suitable for praising and complimenting others. Forum comments contained less explicit network support, which might suggest that the act of replying shows presence. Similar to journal posts and journal comments, here the act of posting may be an indicator of network support. Offering emotional and esteem support more than network might result from an assumption that other members are aware of network support, and need is not explicitly stated. It is possible that users were compelled to offer esteem support when they were more familiar with someone or his or her situation.

One possible cause of pattern differences could be that levels of network or esteem support are correlated with relationship strength. Surprisingly, offered support in journals is different from that in notes, even though their features make them "publicly private." While journals and notes users who communicate with each other might be friends of each other, the longer message format of journals may not be as conducive as notes to maintaining relationships.

Figure 7. Nurturant support offered in messages. FC = forum comments; FP = forum posts; JC = journal comments; JP = journal posts.



Nurturant Support Requested

We identified fewer messages with requested nurturant support than offered nurturant support (Table 7). We calculated the

percentage by dividing the number of messages containing nurturant support by the total number of messages.

Table 7. Number of messages requesting nurturant support. in each format

Support type	Forums		Journals		Notes
	Posts	Comments	Posts	Comments	
Esteem	6	8	21	8	2
Network	8	10	0	0	6
Emotional	29	10	23	7	9
Total	43	28	44	15	17

For requested nurturant support, emotional support was highest in all sets. Esteem support was also frequently requested among all sets of messages but was most noticeable in journal comments, which may indicate a desire on the part of the commenter to help the journal message author feel better about him- or herself. Journal messages did not show network support, possibly because readers were already known friends. In comparison with other formats, in notes and forums requests for network support are quite frequent.

The three patterns that emerged among messages that requested nurturant support were emotional > network > esteem (forum posts, notes); emotional = network > esteem (forum comments); and emotional = esteem (journal posts, journal comments).

The first of these patterns (emotional > network > esteem) appeared in forum posts and notes. The combination of requested supports was an effect of users explicitly stating the type of support they sought, such as an emotional release from thinking about the situation. Notes had more messages requesting network than forum posts, which might emphasize referring to the friendship between the author and receiver. In the forums, more comment messages offered network support than in post messages, which could indicate that people comment on forum threads because they know there is someone else with a similar situation to talk about. The pattern of requested support in notes is most similar to that in forum posts, where network was requested more often than esteem, possibly because they were emphasizing their presence in the online community.

In the second pattern, emotional support occurred in the same number of messages as esteem (forum comments), a demonstration of users showing empathy and appreciation. Perhaps members found talking to each other soothing, especially in the encouragements and time spent chatting with each other. While there was less emotion in forum comments than in forum posts, more esteem and network supports were requested. Perhaps it was easier for members to ask for additional types of support after asking for emotional support at least once before.

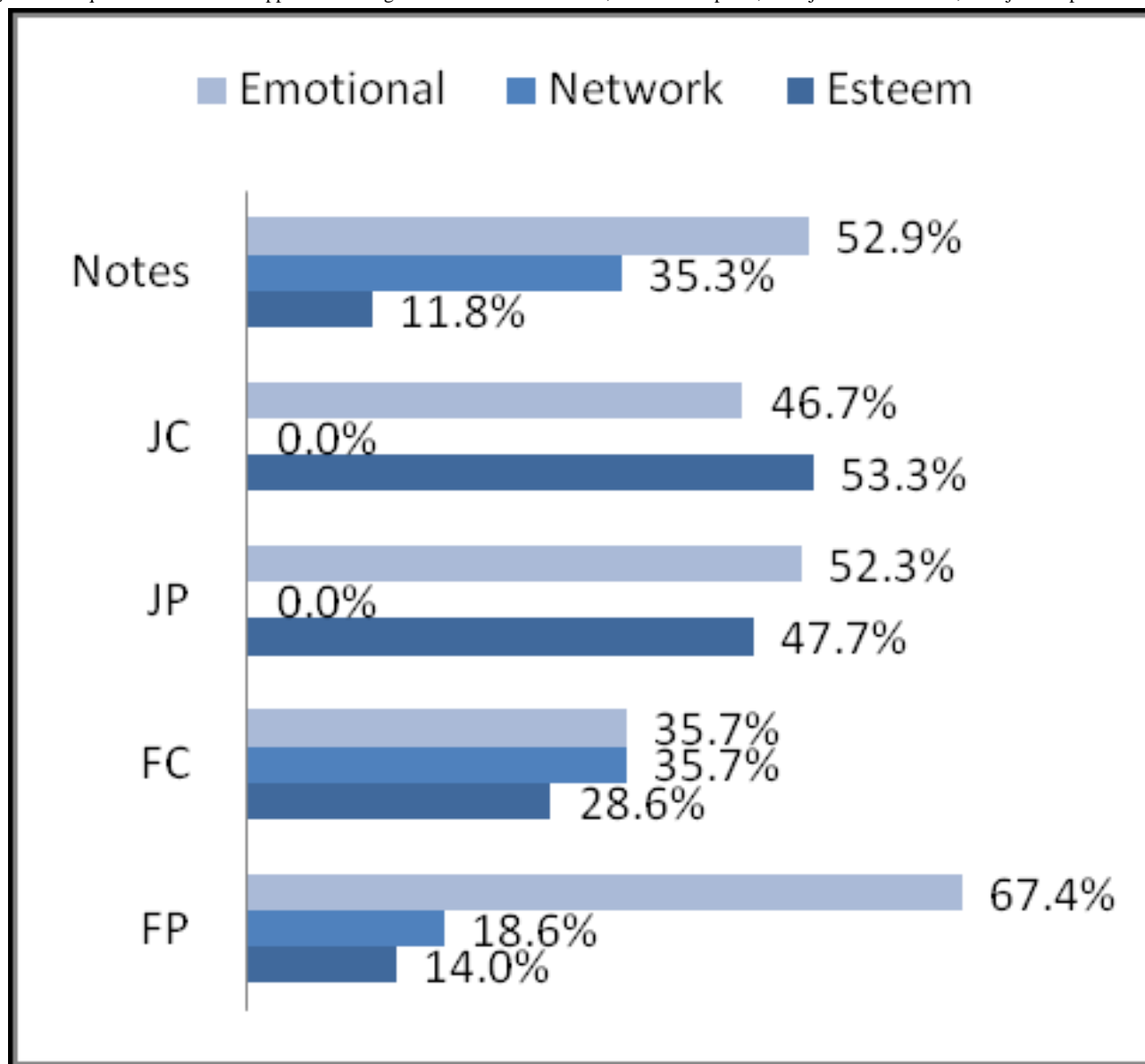
We observed the third pattern of requested nurturant support in journal posts and journal comments, where users were more likely to request only emotional and esteem, rather than network. Because users who write to each other in the journals have a higher likelihood of being friends in the online community or

have a stronger relationship than forum users, they may find it unnecessary to emphasize reminders of network presence, as that might be a purpose more suited for notes. In the case of no network support in a pattern, the architectural features of the tool offered a way out of explicitly stating network support in the note content through the friends feature shown on the profile page. In the journals, users intend to write for themselves or friends, and usually only friends notice new posts and are willing to comment after reading them.

Discussion

This study identified nurturing social support types in user-created messages across three different text-based communication tools (discussion forums, personal journals, and notes) of an online health care social networking site. The content analysis codes came from a literature review and were organized into Cutrona and Suhr's [38] supracategories, and are similar to those from other studies [33,34]. We analyzed the social support in the MedHelp Alcoholism Community most often exchanged by group members during a 3-month period. Interestingly, the nurturant support pattern in forum posts and notes was the same for both the support offered dimension and the support requested dimension, where emotional was greater than both network and esteem support (emotional > network > esteem). These percentages are shown in [Figure 7](#) and [Figure 8](#), and a summary of patterns is explained in [Multimedia Appendix 1](#). The patterns (emotional > esteem > network; emotional = esteem) we found among journal messages (journal posts, journal comments) also group these samples together. Forum comments is the only sample where the pattern of offered support (emotional > esteem > network) was different from requested support (emotional = network > esteem). While the three computer-mediated communication formats of MedHelp have similarities, the differences in architecture appear to have affected the support exchanged.

We found that some formats were more conducive to emotional connecting than others, yet overall each was used for different purposes. This study gathered empirical data of patterns varying nurturing support among the different online tools. This information could be potentially useful to social support scholars and designers of online support groups to understand how software features affect users' behavior. In this section, we explain the results in the context of theories from the related research literature.

Figure 8. Requests for nurturant support in messages. FC = forum comments; FP = forum posts; JC = journal comments; JP = journal posts.

Comparing Patterns From This Study With Related Studies

Similar studies identifying social support in online health support communities found various levels for each nurturant support type [6,19,21,28,32-35]. These studies collected data from settings using bulletin boards or email lists, which is similar to the MedHelp forums. To the best of our knowledge, there have not been any documented cases of health support groups incorporating computer-mediated communication formats similar to MedHelp's journal or note formats, nor any

studies of social support within them. Many studies reporting social network site data are not tailored to a health issue, but rather are general social network sites (eg, Facebook, MySpace, and LinkedIn), where users may already know each other offline. Table 8 summarizes results from other studies identifying social support in health communities in bulletin board or email list styles, rather than social media features such as profile post messages and journal entries. Most studies found that emotional and network support appeared more frequently than esteem support.

Table 8. Nurturant information support interaction patterns in related studies.

Study	Data	Architectural element	Pattern
McCormack [6]	Anorexia bulletin board	Bulletin board	Emotional > network > esteem
Preece [19]	Torn knee ligament, 500 messages, April 1996–April 1997	Email list	Nurturant > information (no distinctions)
Bambina [21]	Support OnLine Cancer Forum, 84 members, 1149 messages (unmoderated)	“an Internet cancer support forum...a virtual space...[to exchange] support;” only requires email address; archives posted online and publicly available	Network > emotional
Braithwaite et al [28]	Support network, 42 users, 1472 messages	“Messages were distributed via email through a nationwide computer BBS network.”	Emotional > esteem > network
Meier et al [31]	10 cancer mailing lists hosted by the Association of Cancer Online Resources (5 months)	Mailing lists	Nurturant > information (no distinctions)
Cunningham et al [32]	Alcoholism, 10 months, 474 posts (moderated)	Bulletin Board	Emotional > esteem
Coursaris and Liu [33]	HIV/AIDS ^a , 5000 messages	Bulletin Board	Emotional > network > esteem
Eichhorn [34]	5 eating disorder message boards, 490 messages	Yahoo Discussion Groups	Emotional > network > esteem
Pfeil and Zaphiris [35]	Depression/seniors	Bulletin Board	Network > emotional > esteem
Selby et al [36]	Smoking cessation	Web assisted tobacco intervention, bulletin board	Esteem > emotional

^a Human immunodeficiency virus/acquired immunodeficiency syndrome.

The members of the Alcoholism Community favored the emotional type of nurturant support, similarly to findings from other health support communities [6,19,28,33,34]. We also noted a substantial amount of informational support in our data in addition to nurturant support. Of all our results, the forum posts sample appears to be most similar to previous findings from bulletin boards (emotional > network > esteem). Notes showed a similar pattern to the forum posts; however, journal messages followed a different pattern with no network support (esteem = emotional). This low amount of network support shows that it is not typical to use journals as a place to indicate a relationship bond between two users, or that network support needs were met simply by participating in the community without the need for explicitly stating requests or offers of network support. In notes, on the other hand, any member can leave a note for a friend and not expect any response. With the journal format, users are given more control over who can read and comment, so in that format initiating an explicit support request is required to receive any comments. Journal post authors are more likely to post a message for themselves. For example, one author wrote in her journal “i doubt if anyone will read this and that is ok as I am writing this more for my benefit then anyone else...” and probably did not expect a reply. In fact, the same post goes on to state “I do not need anyone to tell me how lucky I am, how well off i am, or how good i look. I still feel very fragile and needed to just let someone know that might have a word or encouragement or knidness [sic]...”

Explaining These Patterns

Although offline support is often available from friends and family, e-patients and their caregivers participate in online support groups for the added benefit of specific types of

nurturant support. We found different types of nurturing support in the MedHelp Alcoholism Community of varying patterns across the multiple computer-mediated communication formats. Two theories might be able to explain why the number of support types varies: (1) the purpose of communication affects which format people use to convey and seek help for different types of needs, and (2) the public and private spheres where communication is mediated by particular norms of acceptable behavior affect the content construction for each message.

Limitations

Our results cannot be generalized for two reasons. First, the study setting here is more narrowly defined than in previous studies and we were studying a specific alcoholism community from MedHelp. Second, MedHelp has particular software features for computer-mediated communication that other websites may not have. Alcoholism as a health condition has its own characteristics that can influence attitudes and behaviors. Our hope is to extend more work on each of these reasons to better understand the impact of technology on human interactions.

Purpose of Communication

People use different communication tools for different purposes; for example, some online community members sought information, while others sought compassion and intimacy through conversations [22,50]. This distinction is possibly the result of numerous waves of information needs when recovering from alcoholism [7]. Presenting one's information need(s) to the community may be a way to initiate presence and involvement as a new member, but also for older members to welcome new members. While information is often explicitly

stated within messages, sometimes participation is motivated by other reasons such as relationship maintenance [10,23]. In fact, results from our analysis of notes messages support previous findings from MySpace profile wall posts, which mostly contained short messages to fulfill two purposes: making initial contact and keeping in touch [23]. Because it is so easy to publish information with social media technology, blogs are often used as way to share knowledge and skills, rather than to keep in touch with friends and family [51].

While the purpose of communication can vary across computer-mediated communication formats, it is not the same as the purpose of the community—that is, exchanging support. In this case, the purpose of communication through notes and forum posts (eg, reaching out to others with emotional and network supports) and the purpose of the community (eg, connecting with other patients) overlap. For example, access to other patients' stories on the Internet can be reassuring [52]. In addition, social media make it easier to obtain social feedback and reviews. People who are unsure about medical answers find confidence from social feedback [37]. Health issues trigger anxiety and questions; however, online communication with familiar folk can be soothing, as it might enhance the quality of relationships and improve the psychological well-being of the support seeker [29,53]. For example, most blogs allow readers to leave comments and, in this way, they both generate conversation and encourage collaboration [51]. Users of online support communities often communicate in one-on-one situations or in small groups of 3–5 individuals [29].

Communal Versus Personal Spaces

The different nurturant support patterns in the various computer-mediated communication formats may also be explained by communication theory, which separates communication into that targeted to the public (eg, mass media, advertisements) and private interpersonal communication (eg, email). In recent years, studies of online communication have shown that social media mimic physical spaces by providing online spaces for communication but are also used to exchange private information. The distinction between the traditionally public and private spheres is blurred in online communication [9,17]. We believe that, because privacy can be controlled through notes and journals, we can distinguish these as private spheres, which are more personal, and the forum as a public sphere, where conversations are exchanged in a community setting. While the sphere may be a factor that influences behavior in each format, we believe that the computer-mediated communication format itself is the driving force for different behavior patterns. For example, notes are similar to forum posts, where as forum comments are similar to journal posts and journal comments. It is possible that people are not aware of privacy issues in an online environment or do not find it necessary to control privacy settings or learn about their implications.

In a physical setting, it is easy to perceive the relative privacy of the space. However, in an online environment, the amount of privacy is not transparent, and many private areas become “publicly private.” In this case, perhaps the MedHelp users did not assess the online setting as they would a physical

face-to-face setting. In light of the content observed through this community (eg, blackouts, possible violent episodes), the online setting diminishes the amount of stigma that would be present in face-to-face support.

Emerging Issues: Collaboration and Health Outcomes

In the course of identifying support types, we also found evidence of users in the Alcoholism Community mentioning their health outcomes (eg, quitting drinking) or of collaboration. Although bulletin boards have a less-than-expected effect on behavior change such as smoking cessation, they are often an integral component in Web intervention programs by allowing participants to communicate with each other [54]. Bulletin boards are not suggested as a primary information source or to be appropriate for everybody, as only certain people will voluntarily use them, and they often have a small core set of active users. Participation is also linked to other factors such as an e-patient's phase of quitting (ie, former smokers will have higher participation rate than those prequitting) and the speed of responses to posts.

Social media allow a direct connection between information and the consumer [55]. Social networking enables and facilitates collaboration and collaborative filtering processes. For example, it enables users to see what their peers or others with a predefined relationship (friends, colleagues, fellow patients, etc) are doing; enables automated selection of relevant information (based on what peers are doing and reading on the Web); enables reputation and trust management, accountability, and quality control; and fosters viral dissemination of information and applications (it is this viral marketing aspect that makes Web 2.0 applications so attractive to venture capitalists and public health practitioners alike). Moreover, social networking is a potentially powerful tool to engage users, in that it provides social incentives to enter, update, and manage personal information. Teenagers spend hours keeping their Facebook profile current, constantly updating their status. The same generation of users may possibly turn their attention and energy to similar tools for health.

Future Work

More research to better understand how social support is communicated among computer-mediated communication formats in these groups would be particularly useful to those interested in designing, providing, using, or evaluating online support as an alternative to face-to-face support. Knowing how to develop and sustain an online community is important, as a certain level of active users from peer families and visible signs of their activity are necessary to attract returning users [56]. Potential benefits of social network site participation for e-patients include peer support (availability of an opportunity to receive and offer support), which can be an empowering experience in a customizable setting (eg, revealing limited identity information or restricting the number of friends). Not only that, a personalized space could also help patients open up about their health issues. This could be beneficial to the area of preventive medicine.

Future work that identifies patterns at the message level rather than patterns for each sample would help with understanding

the construction of each message. This study did not account for individual characteristics such as each member's support profile, such as whether he or she is more likely to provide or request support. Future work that identifies whether support is given explicitly to a recipient could be useful.

Another direction for future work is to explore theories on personas at an individual level. For example, the theory of faceted social identity in sociology states that people behave differently with different groups. Users communicate with different categories of relationships online [57]. Users present identities depending on who they are communicating with and seek different types of support. For example, close friends are more likely sources of emotional support. Email is thought of as more personal and private than social network sites, and users are more likely to experience more comfort in using that platform to communicate with close relations [57]. Some users may also be motivated to keep in touch with others, whereas others want to share information [22,50,51]. Age can also be a factor in activeness online.

Yet another research angle for studying patient communities is shared patient data, which is mostly quantitative data (eg, blood pressure) rather than qualitative data such as stories and advice [49,58]. Participants often look for similar patients (by having a medical ailment, hobby, or other thoughts in common) to make their connections. Websites such as PatientsLikeMe "promote data-centered patient conversations." This might be a good direction, as there are many lurkers in support communities [59], who feel that reading community messages is enough to feel a sense of belonging to the collective intelligence. The influence of interpersonal association on personal health records could lead to improved health outcomes as people become more aware of their day-to-day behaviors.

Conclusion

People are drawn to online health support communities because of the availability of tailored information and the opportunities to meet peers who share similar experience(s). Emotional support is an important component of interactions within support groups; however, it varies in exchange across different social media communication formats for reasons such as demographics and formats of communication tools. Users may prefer to ask basic informational questions in the forum because they would

rather have any answer than only the specific answer. Furthermore, users may prefer using the journals to disclose specific information that only a select few people can read. Deeper understanding of these associations would help in designing health-related Internet applications. This in turn can contribute to applications such as online intervention programs.

We found two general patterns in offered nurturant support and three general patterns in requested nurturant support. Offered nurturant support is typically emotional support such as encouragement, and then either esteem (eg, validation) or network (eg, reminders of presence) support. Requested support was also typically emotional support, but with more network than esteem support. There was no mention of network support in the journal posts or comments. We attempted to link theories with results to explain the supportive behaviors. Theoretical perspectives include (1) the purpose of communication, where people use different platforms differently to convey different types of information, and (2) the influence of public and private spheres of communication on the users' behavior. Further research could provide more insight into these two phenomena.

This research offers a novel message regarding the impact computer-mediated communication format has on user interaction patterns in online support communities. It is not clear how people seek or provide social support in an online format, so we explored how the social media platform facilitates the exchange of social support. Social media have more computer-mediated communication features than were present in previously studied software platforms for online communities, in that they give the user a bit more control over whom they share information with, by offering multiple formats for private and public messaging. This research also examined the issue of space preference for privacy and the kinds of support in each format for disclosing information to specific people. We studied how people use social media for nurturant support to have a better understanding of how computer-mediated communication formats can encourage a specific type of social support. For example, people with alcoholism, people who want to quit smoking, or cancer survivors may need more nurturant support. People using other types of health intervention such as weight loss may need more informational support. Therefore, the design and utility of social media can be tailored to the particular purpose.

Conflicts of Interest

None declared

Multimedia Appendix 1

Summary of offered and requested interaction patterns.

[PDF File (Adobe PDF File), 32KB - [jmir_v14i3e54_app1.pdf](#)]

References

1. Høybye MT, Johansen C, Tjørnhøj-Thomsen T. Online interaction. Effects of storytelling in an internet breast cancer support group. *Psycho-Oncology* 2005 Mar;14(3):211-220. [doi: [10.1002/pon.837](#)] [Medline: [15386774](#)]
2. Tufekci Z. Can you see me now? Audience and disclosure management in online social network sites. *Bull Sci Technol Soc* 2007;28(1):20-36. [doi: [10.1177/0270467607311484](#)]

3. Boase J, Horrigan JB, Wellman B, Rainie L. The strength of internet ties. Washington, DC: Pew Internet & American Life Project; 2006 Jan 25. URL: http://www.pewinternet.org/~media/Files/Reports/2006/PIP_Internet_ties.pdf.pdf [accessed 2012-04-04] [WebCite Cache ID 66g3VS3JG]
4. Fox S, Jones S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2009 Jun. URL: http://www.pewinternet.org/~media/Files/Reports/2009/PIP_Health_2009.pdf [accessed 2012-04-04] [WebCite Cache ID 66g3XBGfc]
5. Cohen S. Social relationships and health. *Am Psychol* 2004 Nov;59(8):676-684. [doi: [10.1037/0003-066X.59.8.676](https://doi.org/10.1037/0003-066X.59.8.676)] [Medline: [15554821](https://pubmed.ncbi.nlm.nih.gov/15554821/)]
6. McCormack A. Individuals with eating disorders and the use of online support groups as a form of social support. *Comput Inform Nurs* 2010;28(1):12-19. [doi: [10.1097/NCN.0b013e3181c04b06](https://doi.org/10.1097/NCN.0b013e3181c04b06)] [Medline: [19940616](https://pubmed.ncbi.nlm.nih.gov/19940616/)]
7. Radin PM. "Hello, My Friends": Social Capital and Transformations of Trust in a Grassroots Breast Cancer Website [doctoral thesis]. Seattle, WA: University of Washington; 2001.
8. Wright KB, Bell SB, Wright KB, Bell SB. Health-related Support Groups on the Internet: Linking Empirical Findings to Social Support and Computer-mediated Communication Theory. *J Health Psychol* 2003 Jan;8(1):39-54. [doi: [10.1177/1359105303008001429](https://doi.org/10.1177/1359105303008001429)] [Medline: [22113899](https://pubmed.ncbi.nlm.nih.gov/22113899/)]
9. Papacharissi Z. The virtual geographies of social networks: a comparative analysis of Facebook, LinkedIn and ASmallWorld. *New Media Soc* 2009;11(1-2):199-220. [doi: [10.1177/1461444808099577](https://doi.org/10.1177/1461444808099577)]
10. Ahn YY, Han S, Kwak H, Moon S, Jeong H. Analysis of topological characteristics of huge online social networking services. In: Proceedings. 2007 Presented at: 16th International Conference on World Wide Web; May 8-12, 2007; Banff, AB, Canada p. 835-844. [doi: [10.1145/1242572.1242685](https://doi.org/10.1145/1242572.1242685)]
11. Boyd DM, Ellison NB. Social network sites: definition, history, and scholarship. *J Comput Mediat Commun* 2007;13(1):210-230. [doi: [10.1111/j.1083-6101.2007.00393.x](https://doi.org/10.1111/j.1083-6101.2007.00393.x)]
12. Cunningham JA, Humphreys K, Kypri K, van Mierlo T. Formative evaluation and three-month follow-up of an online personalized assessment feedback intervention for problem drinkers. *J Med Internet Res* 2006;8(2):e5 [FREE Full text] [doi: [10.2196/jmir.8.2.e5](https://doi.org/10.2196/jmir.8.2.e5)] [Medline: [16867968](https://pubmed.ncbi.nlm.nih.gov/16867968/)]
13. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. *J Med Internet Res* 2007;9(2):e10 [FREE Full text] [doi: [10.2196/jmir.9.2.e10](https://doi.org/10.2196/jmir.9.2.e10)] [Medline: [17513281](https://pubmed.ncbi.nlm.nih.gov/17513281/)]
14. Riper H, Kramer J, Keuken M, Smit F, Schippers G, Cuijpers P. Predicting successful treatment outcome of web-based self-help for problem drinkers: secondary analysis from a randomized controlled trial. *J Med Internet Res* 2008;10(4):e46 [FREE Full text] [doi: [10.2196/jmir.1102](https://doi.org/10.2196/jmir.1102)] [Medline: [19033150](https://pubmed.ncbi.nlm.nih.gov/19033150/)]
15. 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](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
16. Preece J. Online Communities: Designing Usability, Supporting Sociability. Chichester: John Wiley; 2000.
17. Clegg Smith KM. "Electronic eavesdropping": the ethical issues involved in conducting a virtual ethnography. In: Johns MD, Chen SLS, Hall GJ, editors. Johns MD, Chen SLS, Hall GJ. Online Social Research: Methods, Issues, & Ethics. New York, NY: Peter Lang Publishers; 2004:223-238.
18. Walstrom M. Seeing and sensing online interaction: an interpretive interactionist approach to USENET support group research. In: Johns MD, Chen SLS, Hall GJ, editors. Johns MD, Chen SLS, Hall GJ. Online Social Research: Methods, Issues, & Ethics. New York, NY: Peter Lang Publishers; 2004:81-97.
19. Preece J. Empathy online. *Virtual Real* 1999;4(1):74-84. [doi: [10.1007/BF01434996](https://doi.org/10.1007/BF01434996)]
20. Beattie MC, Longabaugh R. General and alcohol-specific social support following treatment. *Addict Behav* 1999;24(5):593-606. [Medline: [10574299](https://pubmed.ncbi.nlm.nih.gov/10574299/)]
21. Bambina AD. Online Social Support: The Interplay of Social Networks and Computer-Mediated Communication. Amherst, NY: Cambria Press; 2007.
22. Rau PP, Gao Q, Ding Y. Relationship between the level of intimacy and lurking in online social network services. *Comput Hum Behav* 2008;24(6):2757-2770. [doi: [10.1016/j.chb.2008.04.001](https://doi.org/10.1016/j.chb.2008.04.001)]
23. Thelwall M, Wilkinson D. Public dialogs in social network sites: what is their purpose? *J Am Soc Inf Sci Technol* 2009;61(2):392-404. [doi: [10.1002/asi.21241](https://doi.org/10.1002/asi.21241)]
24. Gilbert E, Karahalios K. Predicting tie strength with social media. 2009 Presented at: 27th International Conference on Human Factors in Computing Systems; Apr 4-9, 2009; Boston, MA, USA. [doi: [10.1145/1518701.1518736](https://doi.org/10.1145/1518701.1518736)]
25. Schaefer C, Coyne JC, Lazarus RS. The health-related functions of social support. *J Behav Med* 1981 Dec;4(4):381-406. [Medline: [7338894](https://pubmed.ncbi.nlm.nih.gov/7338894/)]
26. Thelwall M, Wilkinson D, Uppal S. Data mining emotion in social network communication: gender differences in MySpace. *J Am Soc Inf Sci Technol* 2009;61(1):190-199. [doi: [10.1002/asi.21180](https://doi.org/10.1002/asi.21180)]
27. Haythornthwaite C. Strong, weak, and latent ties and the impact of new media. *Inf Soc* 2002;18(5):385-401. [doi: [10.1080/01972240290108195](https://doi.org/10.1080/01972240290108195)]

28. Braithwaite DO, Waldron VR, Finn J. Communication of social support in computer-mediated groups for people with disabilities. *Health Commun* 1999;11(2):123-151. [doi: [10.1207/s15327027hc1102_2](https://doi.org/10.1207/s15327027hc1102_2)] [Medline: [16370973](https://pubmed.ncbi.nlm.nih.gov/16370973/)]
29. Takahashi Y, Uchida C, Miyaki K, Sakai M, Shimbo T, Nakayama T. Potential benefits and harms of a peer support social network service on the internet for people with depressive tendencies: qualitative content analysis and social network analysis. *J Med Internet Res* 2009;11(3):e29 [FREE Full text] [doi: [10.2196/jmir.1142](https://doi.org/10.2196/jmir.1142)] [Medline: [19632979](https://pubmed.ncbi.nlm.nih.gov/19632979/)]
30. Ancker JS, Carpenter KM, Greene P, Hoffman R, Kukafka R, Marlow LA, et al. Peer-to-peer communication, cancer prevention, and the internet. *J Health Commun* 2009;14 Suppl 1:38-46. [doi: [10.1080/10810730902806760](https://doi.org/10.1080/10810730902806760)] [Medline: [19449267](https://pubmed.ncbi.nlm.nih.gov/19449267/)]
31. Meier A, Lyons EJ, Frydman G, Forlenza M, Rimer BK. How cancer survivors provide support on cancer-related Internet mailing lists. *J Med Internet Res* 2007;9(2):e12 [FREE Full text] [doi: [10.2196/jmir.9.2.e12](https://doi.org/10.2196/jmir.9.2.e12)] [Medline: [17513283](https://pubmed.ncbi.nlm.nih.gov/17513283/)]
32. Cunningham JA, van Mierlo T, Fournier R. An online support group for problem drinkers: AlcoholHelpCenter.net. *Patient Educ Couns* 2008 Feb;70(2):193-198. [doi: [10.1016/j.pec.2007.10.003](https://doi.org/10.1016/j.pec.2007.10.003)] [Medline: [18022340](https://pubmed.ncbi.nlm.nih.gov/18022340/)]
33. Coursaris CK, Liu M. An analysis of social support exchanges in online HIV/AIDS self-help groups. *Comput Hum Behav* 2009;25(4):911-918. [doi: [10.1016/j.chb.2009.03.006](https://doi.org/10.1016/j.chb.2009.03.006)]
34. Eichhorn KC. Soliciting and providing social support over the internet: an investigation of online eating disorder support groups. *J Comput Mediat Commun* 2008;14(1):67-78. [doi: [10.1111/j.1083-6101.2008.01431.x](https://doi.org/10.1111/j.1083-6101.2008.01431.x)]
35. Pfeil U, Zaphiris P. Patterns of empathy in online communication. In: *Proceedings. 2007 Presented at: SIGCHI Conference on Human Factors in Computing Systems; Apr 28-May 3, 2007; San Jose, CA, USA.* [doi: [10.1145/1240624.1240763](https://doi.org/10.1145/1240624.1240763)]
36. Selby P, van Mierlo T, Voci SC, Parent D, Cunningham JA. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. *J Med Internet Res* 2010;12(3):e34 [FREE Full text] [doi: [10.2196/jmir.1340](https://doi.org/10.2196/jmir.1340)] [Medline: [20719739](https://pubmed.ncbi.nlm.nih.gov/20719739/)]
37. Lau AY, Coiera EW. Impact of web searching and social feedback on consumer decision making: a prospective online experiment. *J Med Internet Res* 2008;10(1):e2 [FREE Full text] [doi: [10.2196/jmir.963](https://doi.org/10.2196/jmir.963)] [Medline: [18244893](https://pubmed.ncbi.nlm.nih.gov/18244893/)]
38. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. *Commun Res* 1992;19(2):154-174. [doi: [10.1177/009365092019002002](https://doi.org/10.1177/009365092019002002)]
39. Donath J. Signals in social supernets. *J Comput Mediat Commun* 2007;13(1):231-251. [doi: [10.1111/j.1083-6101.2007.00394.x](https://doi.org/10.1111/j.1083-6101.2007.00394.x)]
40. Meyrowitz J. *No Sense of Place: The Impact of Electronic Media on Social Behavior.* New York, NY: Oxford University Press; 1985.
41. Ellison N, Lampe C, Steinfield C. Social network sites and society: current trends and future possibilities. *Interact Mag* 2009;16(1):6-9. [doi: [10.1145/1456202.1456204](https://doi.org/10.1145/1456202.1456204)]
42. Humphreys L. Mobile social networks and social practice: a case study of Dodgeball. *J Comput Mediat Commun* 2007;13(1):341-360. [doi: [10.1111/j.1083-6101.2007.00399.x](https://doi.org/10.1111/j.1083-6101.2007.00399.x)]
43. Lofland J, Snow DA, Anderson L, Lofland LH. *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis.* Belmont, CA: Thomson Wadsworth; 2006.
44. Med Help International Inc. 2011. MedHelp Alcoholism Community URL: <http://www.medhelp.org/forums/Alcoholism/show/158> [accessed 2012-04-06] [WebCite Cache ID 66j8h2FFz]
45. McKenna KYA, Green AS, Gleason MEJ. Relationship Formation on the Internet: What's the Big Attraction? *J Soc Issues* 2012;58(1):9-31. [doi: [10.1111/1540-4560.00246](https://doi.org/10.1111/1540-4560.00246)]
46. Winzelberg A. The analysis of an electronic support group for individuals with eating disorders. *Comput Hum Behav* 1997;13(3):393-407. [doi: [10.1016/S0747-5632\(97\)00016-2](https://doi.org/10.1016/S0747-5632(97)00016-2)]
47. Chang HJ. Online supportive interactions: using a network approach to examine communication patterns within a psychosis social support group in Taiwan. *J Am Soc Inf Sci Technol* 2009;60(7):1504-1517. [doi: [10.1002/asi.21070](https://doi.org/10.1002/asi.21070)]
48. Civan A, Pratt W. Threading together patient expertise. *AMIA Annu Symp Proc* 2007:140-144. [Medline: [18693814](https://pubmed.ncbi.nlm.nih.gov/18693814/)]
49. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. *J Med Internet Res* 2010;12(2):e19 [FREE Full text] [doi: [10.2196/jmir.1549](https://doi.org/10.2196/jmir.1549)] [Medline: [20542858](https://pubmed.ncbi.nlm.nih.gov/20542858/)]
50. Chuang KY, Yang CC. Helping you to help me: exploring supportive interaction in online health community. *Proc Am Soc Inf Sci Technol* 2010;47(1):1-10. [doi: [10.1002/meet.14504701140](https://doi.org/10.1002/meet.14504701140)]
51. Kovic I, Lulic I, Brumini G. Examining the medical blogosphere: an online survey of medical bloggers. *J Med Internet Res* 2008;10(3):e28 [FREE Full text] [doi: [10.2196/jmir.1118](https://doi.org/10.2196/jmir.1118)] [Medline: [18812312](https://pubmed.ncbi.nlm.nih.gov/18812312/)]
52. Overberg R, Otten W, de Man A, Toussaint P, Westenbrink J, Zwetsloot-Schonk B. How breast cancer patients want to search for and retrieve information from stories of other patients on the internet: an online randomized controlled experiment. *J Med Internet Res* 2010;12(1):e7 [FREE Full text] [doi: [10.2196/jmir.1215](https://doi.org/10.2196/jmir.1215)] [Medline: [20215101](https://pubmed.ncbi.nlm.nih.gov/20215101/)]
53. Bessière K, Pressman S, Kiesler S, Kraut R. Effects of internet use on health and depression: a longitudinal study. *J Med Internet Res* 2010;12(1):e6 [FREE Full text] [doi: [10.2196/jmir.1149](https://doi.org/10.2196/jmir.1149)] [Medline: [20228047](https://pubmed.ncbi.nlm.nih.gov/20228047/)]
54. Stoddard JL, Augustson EM, Moser RP. Effect of adding a virtual community (bulletin board) to smokefree.gov: randomized controlled trial. *J Med Internet Res* 2008;10(5):e53 [FREE Full text] [doi: [10.2196/jmir.1124](https://doi.org/10.2196/jmir.1124)] [Medline: [19097974](https://pubmed.ncbi.nlm.nih.gov/19097974/)]
55. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. *J Med Internet Res* 2008;10(3):e22 [FREE Full text] [doi: [10.2196/jmir.1030](https://doi.org/10.2196/jmir.1030)] [Medline: [18725354](https://pubmed.ncbi.nlm.nih.gov/18725354/)]

56. Nordfeldt S, Hanberger L, Berterö C. Patient and parent views on a Web 2.0 Diabetes Portal--the management tool, the generator, and the gatekeeper: qualitative study. *J Med Internet Res* 2010;12(2):e17 [FREE Full text] [doi: [10.2196/jmir.1267](https://doi.org/10.2196/jmir.1267)] [Medline: [20511179](https://pubmed.ncbi.nlm.nih.gov/20511179/)]
57. Farnham SD, Churchill EF. Faceted identity, faceted lives: social and technical issues with being yourself online. 2011 Presented at: ACM 2011 Conference on Computer Supported Cooperative Work; Mar 19-23, 2011; Hangzhou, China. [doi: [10.1145/1958824.1958880](https://doi.org/10.1145/1958824.1958880)]
58. Frost JH, Massagli MP. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. *J Med Internet Res* 2008;10(3):e15 [FREE Full text] [doi: [10.2196/jmir.1053](https://doi.org/10.2196/jmir.1053)] [Medline: [18504244](https://pubmed.ncbi.nlm.nih.gov/18504244/)]
59. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. *J Med Internet Res* 2008;10(2):e18 [FREE Full text] [doi: [10.2196/jmir.992](https://doi.org/10.2196/jmir.992)] [Medline: [18653442](https://pubmed.ncbi.nlm.nih.gov/18653442/)]

Abbreviations

HIV/AIDS: human immunodeficiency virus/acquired immunodeficiency syndrome

Edited by G Eysenbach; submitted 01.04.11; peer-reviewed by Y Theng, A Lau; comments to author 13.09.11; revised version received 23.01.12; accepted 27.01.12; published 03.05.12.

Please cite as:

Chuang KY, Yang CC

Interaction Patterns of Nurturant Support Exchanged in Online Health Social Networking

J Med Internet Res 2012;14(3):e54

URL: <http://www.jmir.org/2012/3/e54/>

doi: [10.2196/jmir.1824](https://doi.org/10.2196/jmir.1824)

PMID: [22555303](https://pubmed.ncbi.nlm.nih.gov/22555303/)

©Katherine Y Chuang, Christopher C Yang. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 03.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Building a Transnational Biosurveillance Network Using Semantic Web Technologies: Requirements, Design, and Preliminary Evaluation

Douglas Teodoro^{1,2}, BEng; Emilie Pasche^{1,2}, MSc; Julien Gobeill^{2,3}, MSc; Stéphane Emonet^{1,2}, MD; Patrick Ruch³, PhD; Christian Lovis^{1,2}, MPH, MD

¹University Hospitals of Geneva, Geneva, Switzerland

²University of Geneva, Geneva, Switzerland

³University of Applied Sciences Western Switzerland, Geneva, Switzerland

Corresponding Author:

Douglas Teodoro, BEng

University Hospitals of Geneva

Rue Gabrielle-Perret-Gentil 4

Geneva, CH-1201

Switzerland

Phone: 41 (0)22 372 6203

Fax: 41 (0)22 372 6255

Email: douglas.teodoro@hcuge.ch

Abstract

Background: Antimicrobial resistance has reached globally alarming levels and is becoming a major public health threat. Lack of efficacious antimicrobial resistance surveillance systems was identified as one of the causes of increasing resistance, due to the lag time between new resistances and alerts to care providers. Several initiatives to track drug resistance evolution have been developed. However, no effective real-time and source-independent antimicrobial resistance monitoring system is available publicly.

Objective: To design and implement an architecture that can provide real-time and source-independent antimicrobial resistance monitoring to support transnational resistance surveillance. In particular, we investigated the use of a Semantic Web-based model to foster integration and interoperability of interinstitutional and cross-border microbiology laboratory databases.

Methods: Following the agile software development methodology, we derived the main requirements needed for effective antimicrobial resistance monitoring, from which we proposed a decentralized monitoring architecture based on the Semantic Web stack. The architecture uses an ontology-driven approach to promote the integration of a network of sentinel hospitals or laboratories. Local databases are wrapped into semantic data repositories that automatically expose local computing-formalized laboratory information in the Web. A central source mediator, based on local reasoning, coordinates the access to the semantic end points. On the user side, a user-friendly Web interface provides access and graphical visualization to the integrated views.

Results: We designed and implemented the online Antimicrobial Resistance Trend Monitoring System (ARTEMIS) in a pilot network of seven European health care institutions sharing 70+ million triples of information about drug resistance and consumption. Evaluation of the computing performance of the mediator demonstrated that, on average, query response time was a few seconds (mean 4.3, SD 0.1×10^2 seconds). Clinical pertinence assessment showed that resistance trends automatically calculated by ARTEMIS had a strong positive correlation with the European Antimicrobial Resistance Surveillance Network (EARS-Net) ($\rho = .86$, $P < .001$) and the Sentinel Surveillance of Antibiotic Resistance in Switzerland (SEARCH) ($\rho = .84$, $P < .001$) systems. Furthermore, mean resistance rates extracted by ARTEMIS were not significantly different from those of either EARS-Net ($\Delta = \pm 0.130$; 95% confidence interval -0 to 0.030 ; $P < .001$) or SEARCH ($\Delta = \pm 0.042$; 95% confidence interval -0.004 to 0.028 ; $P = .004$).

Conclusions: We introduce a distributed monitoring architecture that can be used to build transnational antimicrobial resistance surveillance networks. Results indicated that the Semantic Web-based approach provided an efficient and reliable solution for development of eHealth architectures that enable online antimicrobial resistance monitoring from heterogeneous data sources. In future, we expect that more health care institutions can join the ARTEMIS network so that it can provide a large European

and wider biosurveillance network that can be used to detect emerging bacterial resistance in a multinational context and support public health actions.

(*J Med Internet Res* 2012;14(3):e73) doi:[10.2196/jmir.2043](https://doi.org/10.2196/jmir.2043)

KEYWORDS

Antimicrobial drug resistance; heterogeneous databases; online information services; surveillance

Introduction

Since their discovery, antibiotics have proved powerful for the control of bacterial infections. However, because of multifactorial causes, especially the widespread use of antibiotics in medicine, animal husbandry, and agriculture, pathogens have developed increasing resistance to many effective drugs [1,2]. The problem of antimicrobial resistance has reached an alarming level, and urgent efforts are needed to avoid regressing to the preantibiotic era [3,4].

In addition to well-known drug resistance cases such as *Pneumococcus* species to penicillin [5-7], outbreaks of new resistant pathogens have become ever more common and have caused many deaths worldwide. Aware of the risks that antimicrobial resistance poses to global public health, the World Health Organization (WHO), among other measures, chose combating antimicrobial resistance as the theme of World Health Day 2011. Lack of effective monitoring systems was identified as an underlying cause of resistance increase, and its improvement is one of the policies the WHO adopted to tackle the problem[8].

Over a decade ago, Monnet et al[9] described and compared the most relevant antimicrobial resistance surveillance systems in Europe. Since then, no new public transnational surveillance initiatives have been developed[10]. Consequently, most projects in use are based either on reporting and manual data acquisition or on outdated information technologies, especially concerning data integration and semantics. Furthermore, no cross-country surveillance system that provides online, direct, and real-time access to antimicrobial resistance information is available. All the systems implemented so far are dependent on delayed data warehouses, usually compiled yearly, which, among other weaknesses, fail to capture antimicrobial resistance outbreaks[10,11]. Finally, these systems do not provide easy ways to export data. Participating institutes have to comply with the surveillance system standards, a labor intensive task, especially for newcomer institutions or newly discovered resistance pathogens [11].

The primary aim of this study was to develop a framework for transnational antimicrobial resistance monitoring, featuring real-time access to laboratory information and being generic with respect to data sources, in order to support multinational resistance surveillance. The secondary aim of the study was to investigate the use of Semantic Web-based architecture in the integration and interoperability of interinstitutional and cross-border databases to support such a framework. To fulfill these aims, we designed the Antimicrobial Resistance Trend Monitoring System (ARTEMIS). ARTEMIS architecture illustrates how Semantic Web technologies can support online

monitoring of antimicrobial resistance trends in heterogeneous networks of health care institutions. It demonstrates how semantically interoperable end points can provide on-demand information on resistance evolution. Furthermore, it describes ways to automate the monitoring process through a state-of-the-art clinical data integration system, which provides mechanisms to adapt to existing electronic health records and laboratory information systems. The architecture is validated according to performance and clinical pertinence.

This paper addresses a large audience, from engineers who have Semantic Web techniques in mind to public health authorities, by showing the results of applying Semantic Web technologies to one of the most crucial current public health challenges: building a global surveillance system for antimicrobial resistances. Here we discuss the technical framework of the project, a technical evaluation, and the quality of the system compared with existing surveillance networks.

Previous European Antimicrobial Resistance Monitoring and Surveillance Initiatives

Several projects have been implemented to provide monitoring and surveillance of antimicrobial resistance evolution in a European context. WHONET was one of the first initiatives to standardize and aggregate results from laboratories in a cross-country environment[12]. Since 1995, the WHO has been developing the WHONET software, in which participating microbiology laboratories present their tests using a specific susceptibility testing terminology defined by the WHO.

The most successful European surveillance project is the European Antimicrobial Resistance Surveillance System[13] developed by the European Centre for Disease Prevention and Control. According to the agency, 900 public health laboratories serving over 1400 hospitals in Europe participate in the network, providing results on a yearly basis. To improve data quality, external control is applied to the susceptibility testing methods used by the participating laboratories. The project has recently evolved into the European Antimicrobial Resistance Surveillance Network (EARS-Net)[14] and will serve as a reference to assess the sampling effectiveness of ARTEMIS.

A few other public initiatives were introduced in parallel. In 1998, the European Society of Biomodulation and Chemotherapy created the European Surveillance of Antibiotic Resistance project[15]. The goal was to establish a representative network of sentinel diagnostic laboratories across Europe to provide antimicrobial resistance monitoring and early detection of new resistant pathogens. In the same year, the US Centers for Disease Control and Prevention launched the International Network for the Study and Prevention of Emerging Antimicrobial Resistance [16] with 79% of participant countries, out of 40, from Europe. The main objective of the project was

to serve as an early warning system for emerging resistant pathogens. Finally, in 1999, the Antimicrobial Resistance Information Bank [17] was derived from the WHONET informal network. Results were reported to the WHO, and an additional external audit quality control was performed on the data. All of these projects have been discontinued, and some were characterized more as a survey than as a surveillance system.

In contrast to the previous initiatives, The Surveillance Network is a corporate-funded surveillance project [18]. It started in 1992 in the United States and later enrolled European laboratories as well. The data extraction and aggregation processes are done by Focus Technologies Inc. (Herndon, VA, USA), the company responsible for the project. Unfortunately, despite having probably the biggest antimicrobial resistance database worldwide, this network provides no antimicrobial resistance information free to the public.

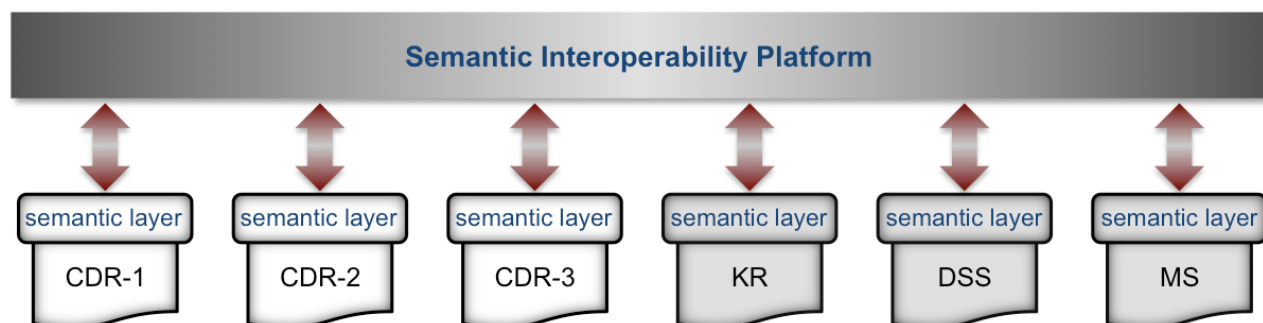
The DebugIT Project

ARTEMIS was developed as part of the Detecting and Eliminating Bacteria Using Information Technology (DebugIT) project, which is funded by the European Union Seventh Framework Programme [19]. DebugIT is a consortium composed of 14 industrial, research, and clinical institutions from nine countries that are collaborating to build a framework for sharing antimicrobial resistance information from clinical

information systems in a Europeanwide context. The project aims to reuse existing clinical data for generating new knowledge to be incorporated in decision support and monitoring engines at the point of care and for developing prevention strategies at policy levels.

The DebugIT architecture (Figure 1) is based on distributed services that exchange information using Semantic Web technologies [20]. The Semantic Web stack provides methods that can contribute to solving technical, syntactic, and semantic differences between disparate data sources [21–24], bringing formal and meaningful representation to data models and sources. First, it presents a standard format to encode information called Resource Description Framework (RDF) [25], which models Web resources in a subject–predicate–object form, a so-called *triple*. This generic model, in contrast to the entity-relationship model used in traditional databases, facilitates the representation of clinical facts to an unconstrained dimension [26]. Second, it has defined the Simple Protocol and RDF Query Language (SPARQL) standard that provides ways to access ubiquitously resources available on the Web [27]. Finally, computer-interpretable ontologies written in the Web Ontology Language [28] bring formal conceptualization to RDF resources, improving the quality of data and fostering interoperability between heterogeneous systems.

Figure 1. Architecture of the Detecting and Eliminating Bacteria Using Information Technology (DebugIT) framework. Components of the architecture, such as the clinical data repository (CDR), knowledge repository (KR), decision support system (DSS), and monitoring system (MS), are interconnected using the HTTP/SPARQL protocol through the Internet bus. Messages are transferred in the RDF format, and ontologies formalize the data model and content.



Methods

Experts from the DebugIT project with different backgrounds, including infectiologists, epidemiologists, computer scientists, knowledge engineers, and eHealth service providers, were involved in the design of ARTEMIS. Over the course of 2 years, we held weekly meetings with these experts to discuss the status of the tasks involved in the system development [29]. In the process, we reviewed the existing distributed integration and interoperable eHealth systems and European antimicrobial resistance monitoring programs. Thereafter, we elaborated the requirements and designed the system model.

Design Requirements

To provide a monitoring system that can be effectively used in the fight against antimicrobial resistance, we derived the following six main requirements based on the published literature and on the expertise of the DebugIT consortium.

The System Shall Provide Online Information

All public European supranational monitoring systems provide resistance information in batch mode—that is, data are collected into batches of laboratory tests and processed periodically, usually on a yearly frequency. While online resistance information is useful on a daily basis at local levels, recent infectious pandemic threats have shown how important this information would be at a multinational level for decision makers. Thus, changing this paradigm to online trends is crucial for antimicrobial resistance surveillance, especially for early warning of emerging resistance trends [10,11].

The System Shall Provide Aggregated Information From Numerous National Sources

Increasing antibiotic resistance is a worldwide public health concern, and for its effective combat, a successful surveillance system has to offer multinational resistance information [30].

The System Shall Not Store Data Centrally

Sharing biomedical data raises several ethical concerns[31]. To comply with international standards on sharing biomedical information, increase the trust of data providers, and encourage collaboration in the surveillance network, central aggregation must be avoided.

The System Shall Implement a Formal and Semantic-Aware Data Model

Most of the available systems do not use formalized biomedical data models, nor computable terminologies and ontologies. As a result, the process of extracting resistance information and data analysis in a heterogeneous environment is done manually or semiautomatically. In addition to the overhead work, the lack of formal conceptualization of the raw laboratory data can have a negative influence on the quality of the data.

The System Shall be High Performing

To be operatively used by health care professionals, whose working environment is recognized to be very time constrained, eHealth systems must provide a fast response time.

The System Shall Provide Reliable Results

Automatic extraction of antimicrobial resistance trends from heterogeneous data sources poses several challenges to accurate data analysis, including concept ambiguity and the common denominator, which can degrade the quality of the examination. However, especially if the system is used by clinicians at the point of care, the accuracy of the results must be equivalent to those obtained by semiautomatic processes, where data cleansing and audit are performed prior to integration and interpretation.

System Model

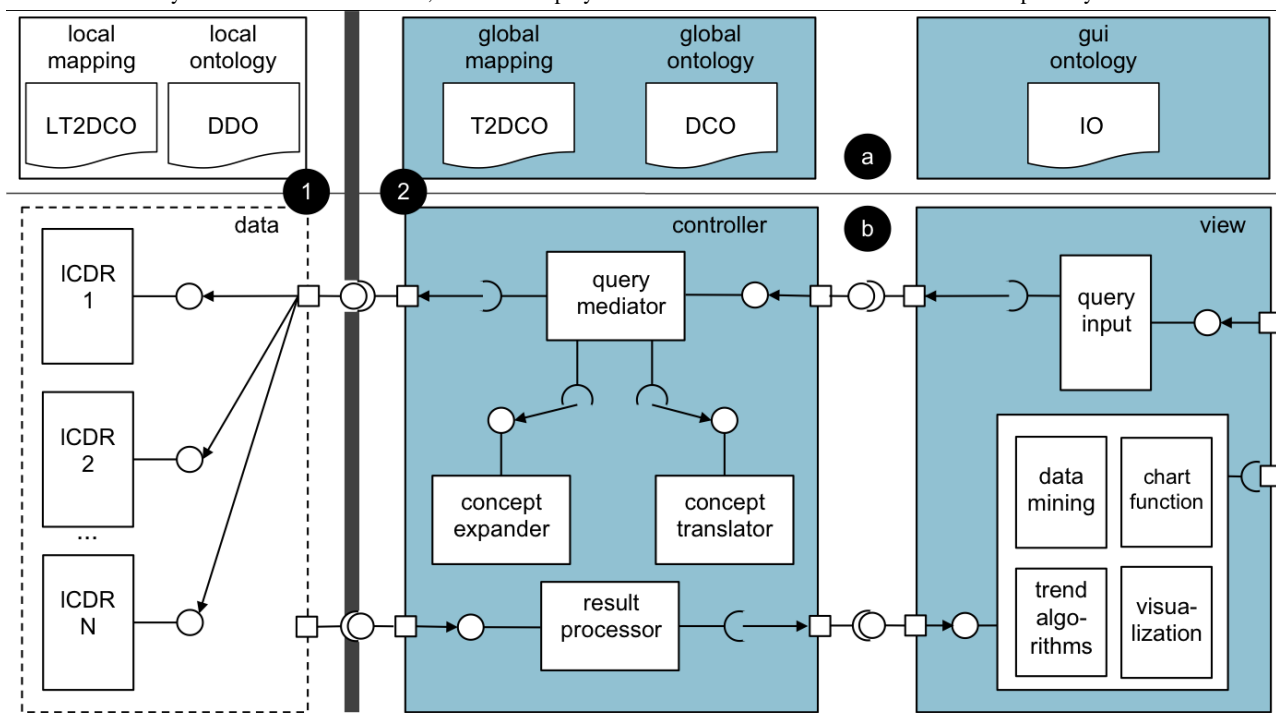
To fulfill the ARTEMIS desiderata, we envisaged the system according to the Semantic Web-complying architecture

presented in [Figure 2](#). The system's semantic interoperability schema is based on an ontology-driven data integration approach[32], where multiple semantically flat local data definition ontologies are mapped to a common domain ontology, the DebugIT Core Ontology [33]. Semantic mappings at local and global levels align concepts from the local ontologies with the domain knowledge.

In the architecture's data model layer, local laboratory databases are connected online to semantic-aware end points, the local clinical data repositories (ICDRs) [34,35]. The ICDRs formalize the local sources and provide a query interface to the controller layer. The semantic mediator, implemented at the controller layer, represents antimicrobial resistance clinical questions as query templates for each end point and coordinates the access to the different sites. It performs the query's data aggregation operations locally to improve query performance and the site's data integration on the fly to avoid central storage. Finally, in the view layer, query templates with parameters extracted from the domain ontology are used to represent antimicrobial resistance clinical questions. As a proof of concept, three initial query templates were proposed by clinicians to be implemented in the system. (1) What is the evolution of resistance to *:antibiotic* of *:bacteria* cultured from *:sample* extracted from *:gender* patients at *:clinical_setting* during period *:begin_date* - *:end_date*? (2) What is the prevalence of *:antibiotic :susceptibility :bacteria* in *:sample* extracted from *:gender* patients at *:clinical_setting* during period *:begin_date* - *:end_date*? (3) What is the rate of *:gender* patients that get *:antibiotic* to treat *:bacteria* infection found in *:sample* at *:clinical_setting* during period *:date_begin* - *:date_end*?

A more detailed description of the system model is given in [Multimedia Appendix 1](#).

Figure 2. Antimicrobial Resistance Trend Monitoring System (ARTEMIS) architecture. (a) Ontology components. Models: data definition ontology (DDO), DebugIT Core Ontology (DCO), and interface ontology (IO). Mappings: local-terminology-to-DCO (LT2DCO) and global-terminology-to-DCO (T2DCO). (b) Run-time business components. (1) Data layer components are deployed within the demilitarized zone of the health care institution. (2) Controller and view layers contain central services, which are deployed in the Internet. ICDR = local clinical data repository.



Participants

To assess ARTEMIS, we connected a network of seven data providers: National Heart Hospital, Sofia, Bulgaria; Les Hôpitaux Universitaires de Genève, Geneva, Switzerland; Georges Pompidou European Hospital, Paris, France; Internetový Prístup Ke Zdravotným Informáciám Pacienta, Prague, Czech Republic; Swedish Intensive Care Registry, Sweden; Athens Chest Hospital “Sotiria”, Athens, Greece; and Universitätsklinikum Freiburg, Freiburg, Germany. [Table 1](#) summarizes antimicrobial resistance-related data shared by these institutions.

We obtained permission to use de-identified data from the ethics committees of the respective participant hospitals. Privacy-sensitive information accessible through the local end points was pseudoanonymized to conform to the European legal and ethical patient data-sharing framework[36]. Data values such as *date of birth* were truncated to the year, and concepts such as *episode of care* (or *encounter*) and *patient identifiers* were encrypted. Furthermore, query templates are pathogen and population centric—that is, the information collected concerns the resistance and treatment of a pathogen population for a given antibiotic in a set of microbiology results. It is therefore not related to a specific patient.

Table 1. Data used in the Antimicrobial Resistance Trend Monitoring System (ARTEMIS).

Data group	Data item	ACH ^a	HEGP ^b	HUG ^c	IZIP ^d	NHH ^e	SIR ^f	UKLFR ^g
Demographics	Age	× ^h	×	×	×	×	×	— ⁱ
	Sex	×	×	×	×	×	×	—
Location	Department	—	×	—	—	—	—	×
Laboratory	Bacteria	×	×	×	—	×	×	×
	Antibiotic	×	×	×	—	×	×	×
	Specimen	×	×	×	—	×	×	×
	S.I.R. ^j	×	×	×	—	×	×	×
Medication	Drug	×	×	×	×	×	—	—
Triples (×10 ⁶)		0.05	25.20	19.87	2.79	0.02	3.81	19.10

^a Athens Chest Hospital “Sotiria.”^b Georges Pompidou European Hospital.^c Les Hôpitaux Universitaires de Genève.^d Internetový Prístup Ke Zdravotným Informáciám Pacienta.^e National Heart Hospital.^f Swedish Intensive Care Registry.^g Universitätsklinikum Freiburg.^h Concept available in the local clinical data repository.ⁱ Concept not available in the local clinical data repository.^j Breakpoint values: susceptible (S), intermediate (I), and resistant (R).

Outcome Measures

The implementation of the functional features defined in the first four design requirements is described at the technical component level using design pattern [37-39] examples. In contrast, for the last two requirements, which can be quantitatively measured, results are presented using efficiency and effectiveness metrics.

Methods for Data Acquisition and Analysis

We measured efficiency using the mediator’s query retrieval time for the three aforementioned query templates. Combinations of pathogens, antibiotics, and sample types were applied to vary the queries and thus avoid database caching effects. Results of the local aggregation mode applied in the query mediator were compared with a central aggregation strategy (baseline).

To assess effectiveness, resistance trends extracted using query template 1 were compared with data from two publicly available surveillance systems: EARS-Net and the Sentinel Surveillance of Antibiotic Resistance in Switzerland (SEARCH). We extracted yearly resistance trends for seven key pathogenic bacteria—*Enterococcus faecalis*, *Enterococcus faecium*, *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Streptococcus pneumoniae*—based on their presence in the three systems. Antibiotics were selected if they were present on both

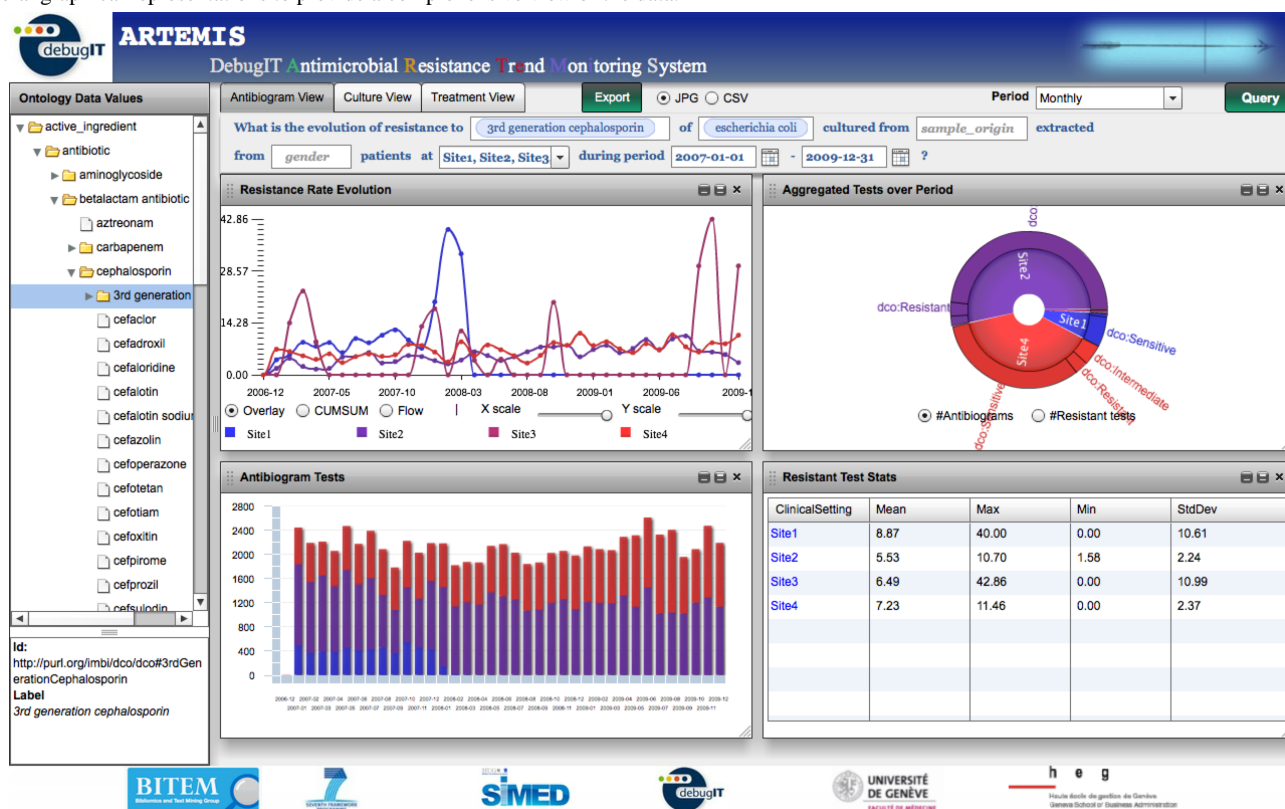
ARTEMIS and the reference system. Resistance rates of the last 4 years (2006 to 2009) available in EARS-Net were used, whereas all years (2008 to 2010) available in SEARCH were taken into account. ARTEMIS data sources that did not contain either more than 1 million triples or data elements to answer the queries were excluded from the analysis, resulting in four sites: Georges Pompidou European Hospital, Les Hôpitaux Universitaires de Genève, Swedish Intensive Care Registry, and Universitätsklinikum Freiburg.

We compared results from Georges Pompidou European Hospital, Swedish Intensive Care Registry, and Universitätsklinikum Freiburg with the resistance rates of their respective EARS-Net countries—France, Sweden, and Germany—and results from Les Hôpitaux Universitaires de Genève with SEARCH. We report correlation and equivalence results using the Spearman rank correlation and the two one-sided convolution[40,41] tests, respectively (see [Multimedia Appendix 2](#)).

Results

ARTEMIS was implemented and deployed in a pilot network of seven European health care institutions sharing 70+ million triples of antimicrobial resistance information. As [Figure 3](#) shows, near real-time resistance trends can be extracted from the distributed network using the system’s Web interface. The tool can be accessed at <http://babar.unige.ch:8080/artemis>.

Figure 3. Antimicrobial Resistance Trend Monitoring System (ARTEMIS) interface. The menu on the left displays the interface ontology concepts, which are used to fill in the template parameters. Each of the view tabs represents a different query template. The data visualization interface displays several graphical representations to provide a comprehensive view of the data.



Model Components

In this section, we present design patterns describing the main functional features of the online distributed monitoring system.

Online Information Provider

Requirement

The system shall provide online information.

Design

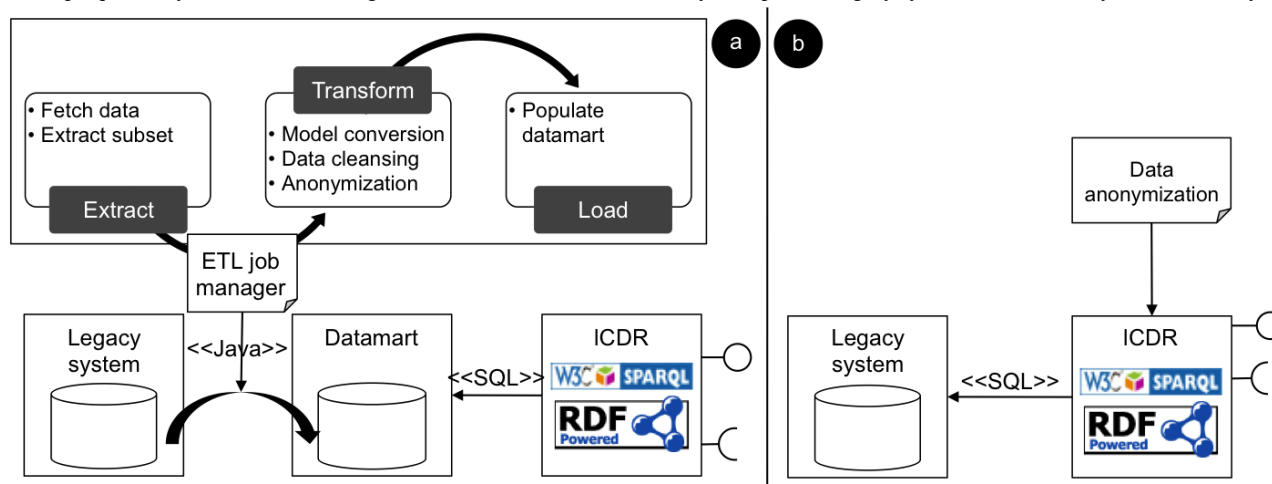
In the architecture presented in Figure 2, local semantic-aware end points, realized by RDF stores, are plugged into the laboratory databases. Thus, microbiology tests are accessible as soon as they are available in the production databases. These

end points are formalized by local ontologies and exposed to the Web so that data are reachable by other parts of the system. In cases where local laboratory databases communicate in the SPARQL protocol, they can be directly connected to the network.

Example

In ARTEMIS, the technical interoperability with the different data sources is provided by D2R [42] engines complemented by site-specific extract, transform, and load processes (Figure 4, part a), which can exploit autocoding methods [43,44]. Alternatively (Figure 4, part b), for cases where there is an accessible production laboratory database, D2R can be plugged directly into the existing system to transform the local data source into a semantic end point (ICDR).

Figure 4. Local clinical data repository (ICDR) deployment and population model. (a) Production data are extracted daily to a local mirror database, which is “sparqlized” by an SQL-to-RDF engine. (b) RDF view is created directly on top of the legacy system. Data are anonymized on the fly.



Distributed Storage

Requirement

The system shall provide aggregated information from numerous international sources.

Design

The technical and semantic heterogeneity within models and concepts from different clinical data sources poses an important barrier for data aggregation and analysis. ARTEMIS architecture relies on a layer of semantically formalized end points, the ICDRs, to solve part of the integration problem. These end points provide a first level of interoperability, modeling the local systems and the data content and providing a common protocol to access data, the SPARQL protocol. The semantic mediator designed in the controller layer builds on top of the ICDR layer and allows the creation of homogeneous aggregated views over the distributed data sources. Thus, the system becomes a grid of semantic-aware sentinels that provide antimicrobial resistance information from heterogeneous supranational data sources.

Example

In ARTEMIS, the ICDRs are provided by RDF-like stores to create the first semantic layer on top of the local databases. The data definition ontologies formalize the local end points and expose linkable data on the Web. The Jena Framework is used for querying the remote ICDRs and for reasoning over the RDF models.

Institutional Autonomy

Requirement

The system shall not store data centrally.

Design

ARTEMIS changes the centralized integration paradigm used in antimicrobial resistance surveillance. Unlike other systems [9-11], its distributed architecture does not require centralization of microbiology test results. At query time, a global aggregated view on the local end points is created by the

semantic mediator, solving the problem of interoperability while avoiding a central repository, which would violate the project's legal requirements. Additionally, since there is no need to move data across the health care border, this design gives full control to participating sites, allowing them to stop sharing data at any moment. Further, no historical information for the respective site is kept on the system.

Example

In the model-view-control pattern [37] presented in Figure 2, persistent data stores are deployed only within the demilitarized zone of the data providers. The central mediator process and aggregates query constraints locally. In this configuration, there is no need to move datasets with information at the patient level out of the institutional borders. Only aggregated population data are retrieved at query time. Furthermore, institutions can stop sharing data at any moment by shutting down the ICDR server. This change is automatically reflected in ARTEMIS, which will not be able to retrieve any data from the respective data source; other sources remain seamlessly reachable.

Knowledge Representation

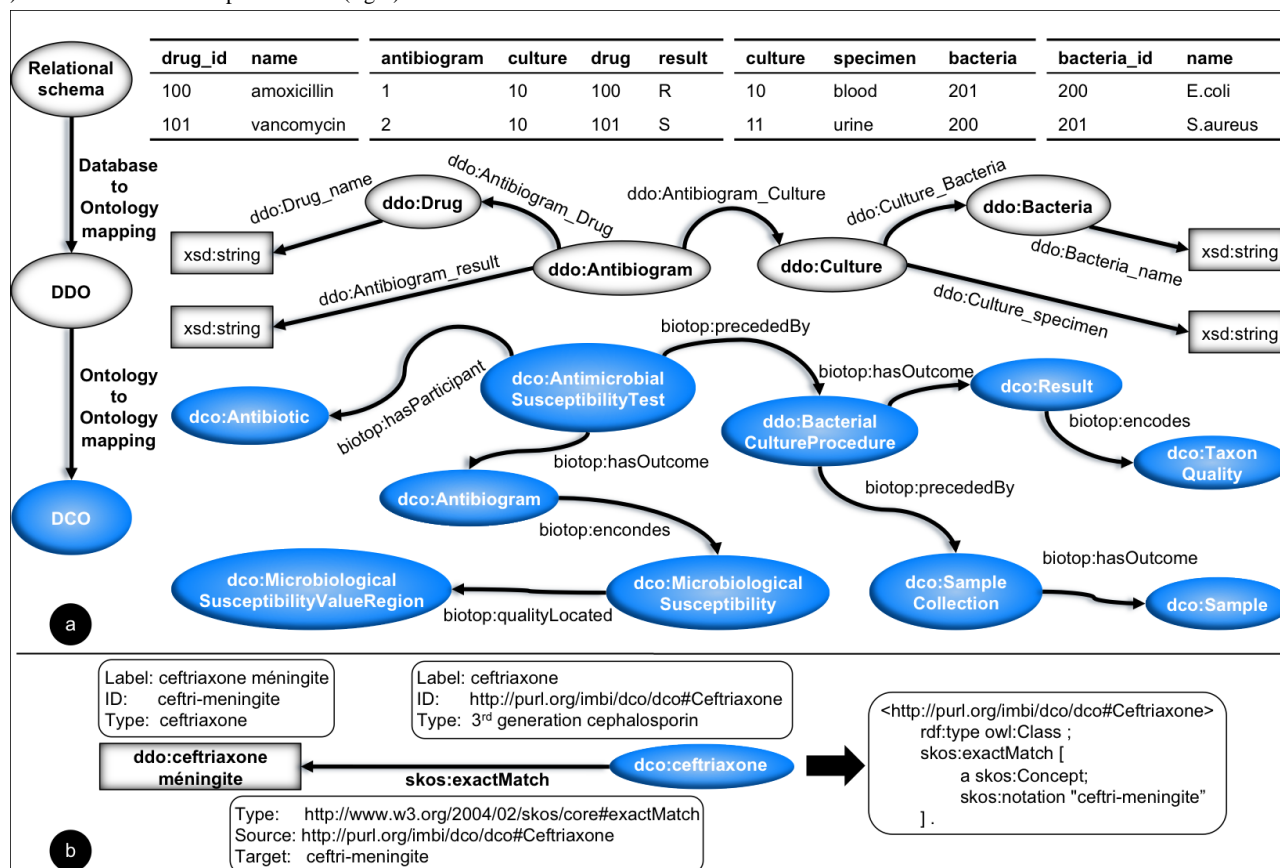
Requirement

The system shall implement a formal and semantic-aware data model.

Design

In a multinational environment, the contents of electronic health records and laboratory information systems are expressed in several languages and different terminologies. Additionally, spelling mistakes and abbreviations are common in concept definitions. These ambiguities reduce the quality of the statistical analysis. To have unified semantics across the different data sources, in ARTEMIS's knowledge model (Figure 5), concepts are represented using a formal language (RDF/OWL). Further, they are aligned into common syntaxes defined by biomedical terminologies. Finally, to have a common meaning across the whole system, these formally represented terminologies are mapped to a shared domain ontology.

Figure 5. The hybrid ontology-driven interoperability mapping model. White elements represent local-level concepts and blue elements represent shared knowledge. (a) Local entity-relationship schemata are formalized by the data definition ontologies (DDOs). Mappings between DDO data elements and DebugIT Core Ontology (DCO) link local concepts to the global knowledge. (b) Example of a semantic mapping: concept map diagram (left) and RDF/Notation3 representation (right).



Example

In ARTEMIS, standard terminologies such as the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT), the WHO's Anatomical Therapeutic Chemical (WHO-ATC) classification system, and the Universal Protein Resource (UniProt/NEWT) are mapped to DebugIT Core Ontology using the Simple Knowledge Organization System (SKOS) ontology [45] and Notation3 rules (Figure 5b). If local concepts are not already defined using these terminologies, they are normalized against them using automatic classification tools [43,44]. Alternatively, local concepts represented in the SKOS notation can be directly mapped to DebugIT Core Ontology.

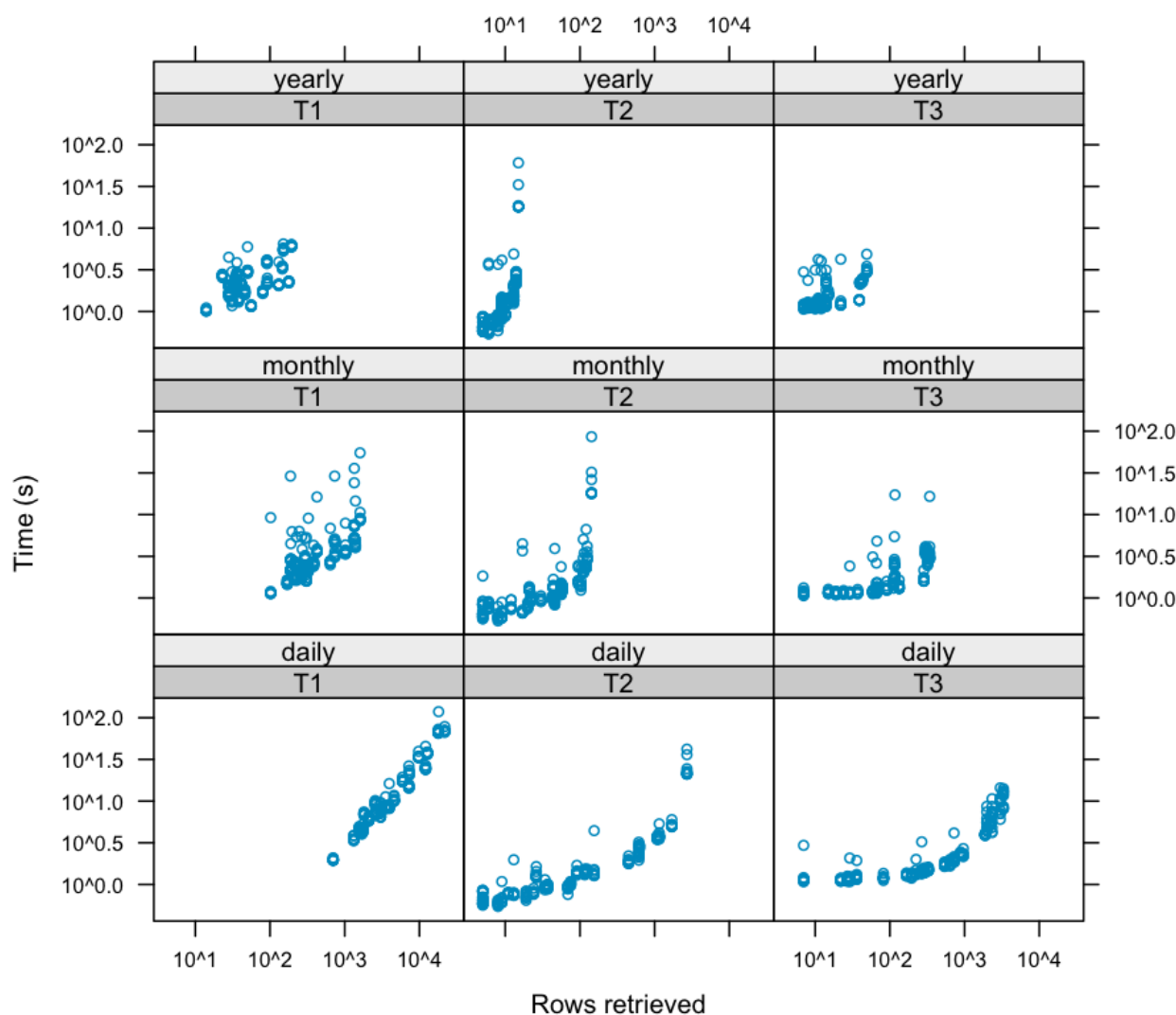
Performance Requirement

We assessed the mediator's SPARQL query performance for query templates 1, 2, and 3. A query mix composed of 225 unique queries, spanning 4 years in daily, monthly, and yearly periods, were used. Each query mix was submitted 10 times against the seven end points. Table 2 summarizes the results. The mean query response time was 4.3 (SD 0.1×10^2) seconds. Comparing the results with a different aggregation strategy, based on central reasoning, the average retrieval time increased almost 30-fold (mean 130.5, SD 0.1×10^3 seconds).

Figure 6 shows how the response time of ARTEMIS queries varied with the number of rows retrieved for different query templates and aggregation periods. Indeed, the response time is highly correlated with the number of rows retrieved ($\rho = .81$, $P < .001$).

Table 2. Arithmetic (t_a) and geometric (t_g) mean (SD) execution times for the two query mediation strategies: local (Antimicrobial Resistance Trend Monitoring System [ARTEMIS]) versus central (baseline) reasoning.

Template	Number of distinct queries	ARTEMIS t_a (seconds)	ARTEMIS t_g (seconds)	Baseline t_a (seconds)	Baseline t_g (seconds)
1	75	8.4 (0.1×10^2)	4.2 (0.1)	311.0 (0.9×10^3)	308.3 (0.1)
2	75	2.3 (0.6×10)	1.3 (0.1)	74.7 (0.6×10^2)	72.1 (0.1)
3	75	2.0 (0.2×10)	1.7 (0.1)	5.9 (0.8×10)	2.7 (0.1)
All	225	4.3 (0.1×10^2)	2.1 (0.1)	130.5 (0.1×10^3)	39.2 (0.1)

Figure 6. Query performance. Response time and rows retrieved by template (1-3) and aggregation period. As the number of rows retrieved increases, the response time tends also to increase.

Result Reliability Requirement

Following the data selection criterion, we created 221 queries for EARS-Net and 153 for SEARCH based on template 1. Table 3 shows the geometric mean resistance rates extracted from the three systems. The results yielded a strong positive correlation coefficient between ARTEMIS and both EARS-Net ($\rho = .86$, $P < .001$) and SEARCH ($\rho = .84$, $P < .001$) reference systems.

The within-country geometric standard deviation of EARS-Net was $\sigma_{\text{ears}} = 0.130$. This value was extrapolated to the similarity region Δ ($\Delta = \sigma_{\text{ears}}$) of the two one-sided convolution test. Figure 7 (part a, all results and part b, without outliers) presents the correlation between the two systems, and Figure 7c shows the regions of similarity. The confidence interval (CI) lies in the region of similarity (95% CI 0–0.030; $P < .001$), confirming

the equivalence between the ARTEMIS and EARS-Net resistance rates. Similarly, for SEARCH, the Swiss region's geometric standard deviation was $\sigma_{\text{search}} = 0.042$, indicating a small susceptibility rate variation in the different regions. In this scenario, the results of ARTEMIS (Figure 8, part a) cannot

be considered equivalent to SEARCH (95% CI 0–0.052; $P = .18$). However, removing outliers—that is, those results that fall within a difference in resistance rate bigger than $3\sigma_{\text{search}}$ (Figure 8, part b)—also leads to an equivalent outcome (95% CI –0.004 to 0.028; $P = .004$).

Table 3. Resistance rate geometric mean (SD) and correlation results.

Number of queries	Resistance rate			ρ	P value
	EARS-Net ^a	SEARCH ^b	ARTEMIS ^c		
221	0.032 (0.002×10 ²)	NA ^d	0.038 (0.002×10 ²)	.86	<.001
153	NA	0.042 (0.001×10 ²)	0.053 (0.002×10 ²)	.84	<.001

^a European Antimicrobial Resistance Surveillance Network.

^b Sentinel Surveillance of Antibiotic Resistance in Switzerland.

^c Antimicrobial Resistance Trend Monitoring System.

^d Not applicable.

Figure 7. Antimicrobial Resistance Trend Monitoring System (ARTEMIS) vs European Antimicrobial Resistance Surveillance Network (EARS-Net). (a) Resistance rates ($n = 221$). Black line indicates an exact match (100% equivalence). Gray line indicates best fit. Gray dashed lines indicate $\Delta = \pm 0.130$. (b) Resistance rates without outliers ($n = 213$). (c) Gray vertical dashed lines indicate similarity region Δ . Gray horizontal bars indicate two one-sided convolution confidence interval (CI). 95% CI^a 0–0.030 ($P < .001$); 95% CI^b 0.002–0.026 ($P < .001$).

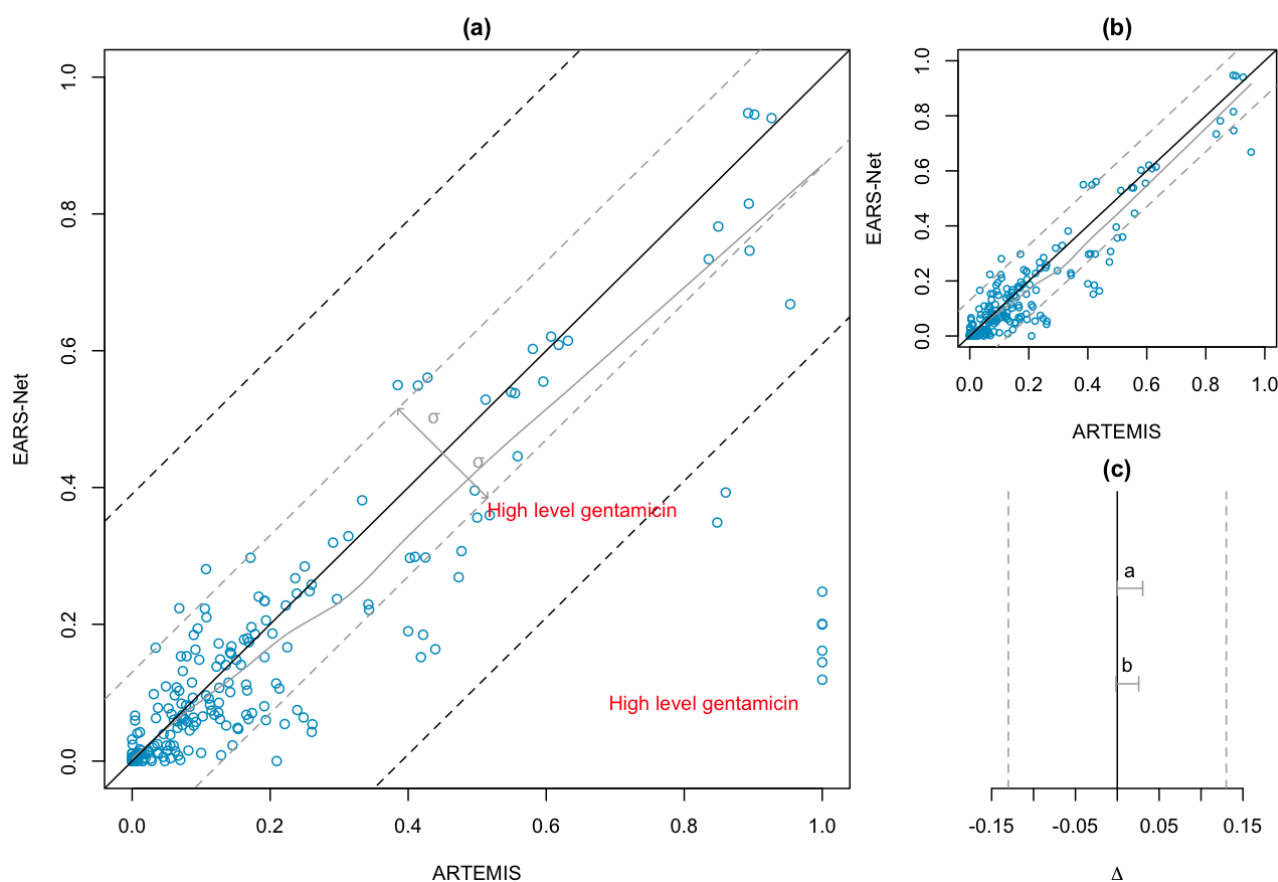
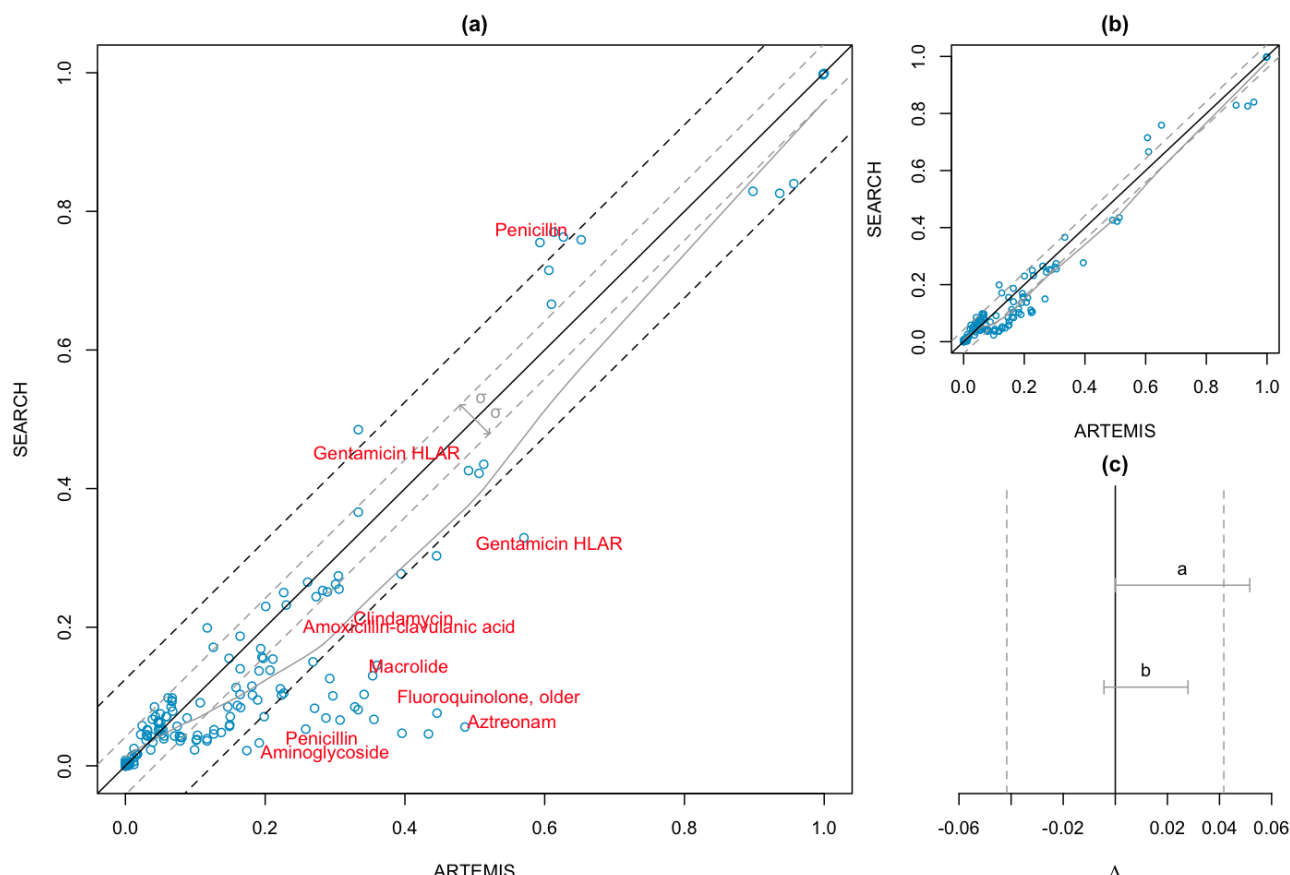


Figure 8. Antimicrobial Resistance Trend Monitoring System (ARTEMIS) vs Sentinel Surveillance of Antibiotic Resistance in Switzerland (SEARCH). (a) Resistance rates ($n = 153$). Black line indicates exact match (100% equivalence). Gray line indicates best fit. Gray dashed lines indicate $\Delta = \pm 0.042$. (b) Resistance rates without outliers ($n = 143$). (c) Gray vertical dashed lines indicate similarity region Δ . Gray horizontal bars indicate two one-sided convolution confidence interval (CI). 95% CI^a 0–0.052 ($P = .17$); 95% CI^b –0.004 to 0.028 ($P = .004$).



Discussion

In this paper, we present an online and source-independent architecture that enables monitoring of multinational microbiology databases. The system was implemented and deployed in a pilot surveillance network distributed across Europe. From the results, one can see that Semantic Web-based architectures such as that of ARTEMIS are suitable to automate the integration and interoperability of distributed microbiology laboratory data sources. Therefore, it can be used to enable automatic access to antimicrobial resistance information in a transnational context and foster real-time multinational biosurveillance. The architecture is able to interoperate heterogeneous networks via the use of semantic maps that account for local specificity. The data integration process is performed on the fly using standard end points powered with RDF/SPARQL communication, which are mediated by a central engine. The local end points are directly connected to the laboratories' databases and as such are able to provide (near) real-time resistance information, while avoiding centralization of the data.

System Architecture

The data integration architecture proposed in ARTEMIS is distinct from existing antimicrobial resistance surveillance systems [14,17,18], as it implements a loosely coupled data federation design[46], which is realized by formalization of the

data sources and data semantics. Thus, the data layer is detached from the central system, which avoids central storage and guarantees to care providers full control over the local information. Moreover, online semantic data repositories automate access to local antimicrobial resistance databases, allowing the system to retrieve near real-time antimicrobial resistance trends. Therefore, emerging and outbreak resistances can be easily monitored on a multinational scale. Finally, instead of predetermined and statically monitored bacteria–antibiotic pairs, the architecture introduced here facilitates the expansion of the concept coverage, making the process of tracking resistance of new antibiotics and bacteria trivial. Since concepts are fully formalized by ontologies through the whole architecture, to add a new item to be monitored it is only necessary to create the respective class in the domain ontology and represent it in the semantic mappings (global, local, or both). Thus, it will be automatically reflected in the user interface, including past occurrences of the given class in microbiology tests.

ARTEMIS uses open Semantic Web technologies to provide technical and semantic interoperability. Semantic data sources create a common technical layer over the local microbiology databases, which can be accessed through a standard query protocol (SPARQL). Since local end points are fully formalized and accessible through the Web, they can be linked to external Web resources, such as the Linked Life Data[47], or reused in

other clinical research projects to leverage knowledge on infectious diseases by combining different sources of information. Another benefit of using ontologies to represent data is the hierarchical structure, which allows higher-level representation of concepts. Therefore, the system can handle complex queries expressed at group levels allowing, for example, automatic clustering of antibiotic classes such third-generation cephalosporin or bacteria families such as Enterobacteriaceae.

Finally, the powerful query interface allied with the availability of near real-time results makes ARTEMIS not only useful to bodies concerned with supranational resistance but also potentially beneficial to local needs, especially if connected to online prescribing systems for empirical treatments. In addition, this local application might facilitate the maintenance of the system by health care institutions. As Goble and Stevens discussed [46], data integration systems tend to become “data mortuaries” once the research funding ends. Local appeal can possibly help to change this pattern.

Performance

All SPARQL performance benchmarks presented in the literature are focused on local single-source servers[48]. Thus, they are not adequate to assess the performance of data integration systems. Hence, the ARTEMIS semantic mediator was compared with a standard approach of retrieving and aggregating centrally. As Table 2 shows, the push-down procedure has reduced the retrieval time by 30-fold (19-fold considering the geometric mean). Indeed, as Figure 6 shows, in a distributed system, response time is nearly linearly correlated ($p = .81$, $P < .001$) with the amount of data retrieved. Thus, local reasoning is crucial for systems that require fast response time.

The preference for an SQL-to-RDF engine [42] instead of a native RDF triple store to formalize local data sources was due to scalability issues. As Schmidt et al [49] noted, native RDF triple stores can hardly be scaled to answer queries when their size is bigger than a few million triples. At the mediation level, the use of a push-down approach while performing aggregation has proved efficient. The average query response was in the order of a few seconds (mean 4.3, SD 0.1×10^2 seconds), which could contribute to the adoption of the system by clinicians, who consider a good response time an important requirement in the system design [50].

Comparison with Existing Systems

Existing surveillance systems normally use semiautomatic methods to extract antimicrobial resistance rates[51]. Validation and cleansing steps are taken by experts before statistical analysis. In ARTEMIS, this process is fully automated and, as such, errors can be introduced. To validate ARTEMIS resistance trends, we compared antimicrobial resistance rates with European and national reference systems. The results indicated a strong positive correlation between the susceptibility test outcomes. We carried out a second evaluation based on equivalence tests to confirm the trustworthiness of the results. The tests showed that at the limit of 3σ ARTEMIS trends are deemed equivalent to both EARS-Net and SEARCH.

A difference in concept definition between ARTEMIS and the reference systems negatively affected the results. The majority of outliers (18 out of 33) presented in Figure 7a and Figure 8a were caused by semantic ambiguities between concepts. For example, in ARTEMIS, antibiotic definition follows the WHO-ATC classification system terminology, which does not define a single antibiotic concept for penicillin but rather classes including several antibiotics based on penicillin. In SEARCH, this concept is defined as an antibiotic agent. Analogously, the gentamicin definition, which is not related to concentration in ARTEMIS, is defined as *Gentamicin HLAR* in SEARCH and *High level gentamicin* in EARS-Net. These issues were not accentuated in the comparison with EARS-Net because, as expected, the region of similarity was wider than that of SEARCH, which considers only within-country variations. Adoption of standard and formalized terminologies in the eHealth care field and a more dynamic evolution of terminological resources so that they can cover operational needs are part of the semantic solution.

Finally, in statistical analysis, care should be taken with duplicate tests. If all apparent duplicates are ignored indiscriminately, information may be omitted, such as nosocomial infection, whereas inclusion of all tests may skew the results, usually toward augmentation of resistance [10]. In the reference systems, duplicate tests are manually removed. In ARTEMIS, biases were automatically minimized by considering only the unique tests within an episode of care.

Limitations

In an ontology-based integration system, automatic mapping from global to local ontologies using first-order logic reasoners creates logical inconsistencies because knowledge from the various local ontologies cannot be completely reconciled in the global model[52]. For example, if at site 1 vancomycin-resistant *Enterococcus* is prevalent, this fact is not necessary true for all other sites. A solution, as implemented in ARTEMIS, is to create query templates over the local ontologies. However, as the system expands to a large number of clinical providers, this approach may prove difficult to maintain, since query templates must be defined centrally for each new data source. Nevertheless, this limitation could be easily overcome if local sources provided a datamart with a common data model as proposed in Figure 4a.

Aligning multinational microbiology laboratory results presents several issues. For example, it has been shown[16] that, for a given sample test, independent laboratories will present different outcomes. Differences in susceptibility breakpoint across countries is also a complex issue involving standardization of antibiogram methodologies. Additionally, results of second-line antibiotics tend to present bias toward resistance, since they are normally tested when isolates show resistance to first-line drugs[10]. The methodology proposed here cannot solve most of the intrinsic divergence between different laboratory procedures. Regardless, ARTEMIS does not aim to tackle these issues but rather to promote access to distributed antimicrobial resistance information as soon as data are available in a formalized and semantically defined way.

Conclusions

We designed, implemented, and deployed the ARTEMIS architecture in a small-scale biosurveillance network of European hospitals. Results indicate that the distributed monitoring architecture introduced here can potentially be used to build transnational antimicrobial resistance surveillance networks. The architecture proved to be efficient and reliable, while complying with local legal and regulatory frameworks.

The Semantic Web-based approach proved to be an effective solution for development of eHealth architectures that enable online antimicrobial resistance monitoring from heterogeneous data sources. In the future, we plan to investigate local model mediation, paving the way to a more easily maintainable system. We expect that new health care institutions can join the network so that it can provide clinicians and decision makers with a missing tool to tackle the growing threat of rising emergent infectious diseases and antibiotic resistance patterns.

Acknowledgments

This work is funded by the DebugIT project of the European Union Seventh Framework Programme grant agreement ICT-2007.5.2-217139.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Architecture design.

[PDF File (Adobe PDF File), 355KB - [jmir_v14i3e73_app1.pdf](#)]

Multimedia Appendix 2

Equivalence test.

[PDF File (Adobe PDF File), 56KB - [jmir_v14i3e73_app2.pdf](#)]

References

1. Adam D. Global antibiotic resistance in *Streptococcus pneumoniae*. *J Antimicrob Chemother* 2002 Jul;50 Suppl:1-5 [FREE Full text] [Medline: [12077153](#)]
2. Sarmah AK, Meyer MT, Boxall AB. A global perspective on the use, sales, exposure pathways, occurrence, fate and effects of veterinary antibiotics (VAs) in the environment. *Chemosphere* 2006 Oct;65(5):725-759. [doi: [10.1016/j.chemosphere.2006.03.026](#)] [Medline: [16677683](#)]
3. World Health Organization. 2001. WHO Global Strategy for Containment of Antimicrobial Resistance URL: http://whqlibdoc.who.int/hq/2001/WHO_CDS_CSR_DRS_2001.2.pdf [accessed 2012-01-02] [WebCite Cache ID 64Oez8NaK]
4. Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, et al. Bad bugs, no drugs: no ESCAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis* 2009 Jan 1;48(1):1-12 [FREE Full text] [doi: [10.1086/595011](#)] [Medline: [19035777](#)]
5. Pittet D, Safran E, Harbarth S, Borst F, Copin P, Rohner P, et al. Automatic alerts for methicillin-resistant *Staphylococcus aureus* surveillance and control: role of a hospital information system. *Infect Control Hosp Epidemiol* 1996 Aug;17(8):496-502. [Medline: [8875292](#)]
6. Jacobs MR. Worldwide trends in antimicrobial resistance among common respiratory tract pathogens in children. *Pediatr Infect Dis J* 2003 Aug;22(8 Suppl):S109-S119. [Medline: [14566997](#)]
7. Devaux I, Manissero D, Fernandez de la Hoz K, Kremer K, van Soolingen D, EuroTB network. Surveillance of extensively drug-resistant tuberculosis in Europe, 2003-2007. *Euro Surveill* 2010 Mar 18;15(11) [FREE Full text] [Medline: [20338144](#)]
8. World Health Organization. 2011 Feb. Antimicrobial Resistance Fact Sheet no 194 URL: <http://www.who.int/mediacentre/factsheets/fs194/en/> [accessed 2012-01-02] [WebCite Cache ID 64OgMkz8X]
9. Monnet DL. Toward multinational antimicrobial resistance surveillance systems in Europe. *Int J Antimicrob Agents* 2000 Jul;15(2):91-101. [Medline: [10854804](#)]
10. Giske CG, Cornaglia G, ESCMID Study Group on Antimicrobial Resistance Surveillance (ESGARS). Supranational surveillance of antimicrobial resistance: The legacy of the last decade and proposals for the future. *Drug Resist Updat* 2010 Oct;13(4-5):93-98. [doi: [10.1016/j.drug.2010.08.002](#)] [Medline: [20864384](#)]
11. O'Brien TF, Stelling J. Integrated Multilevel Surveillance of the World's Infecting Microbes and Their Resistance to Antimicrobial Agents. *Clin Microbiol Rev* 2011 Apr;24(2):281-295 [FREE Full text] [doi: [10.1128/CMR.00021-10](#)] [Medline: [21482726](#)]

12. Stelling JM, O'Brien TF. Surveillance of antimicrobial resistance: the WHONET program. *Clin Infect Dis* 1997 Jan;24 Suppl 1:S157-S168 [FREE Full text] [Medline: 8994799]
13. Bronzwaer SL, Goettsch W, Olsson-Liljequist B, Wale MC, Vatopoulos AC, Sprenger MJ. European Antimicrobial Resistance Surveillance System (EARSS): objectives and organisation. *Euro Surveill* 1999 Apr;4(4):41-44 [FREE Full text] [Medline: 12631906]
14. European Centre for Disease Prevention and Control. 2011. European Antimicrobial Resistance Surveillance Network (EARS-Net) URL: <http://www.ecdc.europa.eu/en/activities/surveillance/EARS-Net/Pages/index.aspx> [accessed 2012-01-02] [WebCite Cache ID 64Oj5wa68]
15. European Society of Biomodulation and Chemotherapy Datacenter Munich. 1999. European Surveillance of Antibiotic Resistance (ESAR) URL: http://www.esbic.de/esbic/ind_esar.htm [accessed 2012-01-02] [WebCite Cache ID 64OjLRqte]
16. Richet HM, Mohammed J, McDonald LC, Jarvis WR. Building communication networks: international network for the study and prevention of emerging antimicrobial resistance. *Emerg Infect Dis* 2001 Apr;7(2):319-322 [FREE Full text] [Medline: 11294732]
17. Anonymous. The WHO Antimicrobial Resistance Information Bank. *WHO Drug Inf* 1999;13(4).
18. Karlowsky JA, Kelly LJ, Thornsberry C, Jones ME, Evangelista AT, Critchley IA, et al. Susceptibility to fluoroquinolones among commonly isolated Gram-negative bacilli in 2000: TRUST and TSN data for the United States. *Tracking Resistance in the United States Today. The Surveillance Network. Int J Antimicrob Agents* 2002 Jan;19(1):21-31. [Medline: 11814764]
19. Lovis C, Colaert D, Stroetmann VN. DebugIT for patient safety - improving the treatment with antibiotics through multimedia data mining of heterogeneous clinical data. *Stud Health Technol Inform* 2008;136:641-646. [Medline: 18487803]
20. Berners-Lee T, Hendler J, Lassila O. The semantic web. *Sci Am* 2001;May:29-37.
21. Sahoo SS, Bodenreider O, Rutter JL, Skinner KJ, Sheth AP. An ontology-driven semantic mashup of gene and biological pathway information: application to the domain of nicotine dependence. *J Biomed Inform* 2008 Oct;41(5):752-765. [doi: 10.1016/j.jbi.2008.02.006] [Medline: 18395495]
22. Belleau F, Nolin MA, Tourigny N, Rigault P, Morissette J. Bio2RDF: towards a mashup to build bioinformatics knowledge systems. *J Biomed Inform* 2008 Oct;41(5):706-716. [doi: 10.1016/j.jbi.2008.03.004] [Medline: 18472304]
23. Chen B, Dong X, Jiao D, Wang H, Zhu Q, Ding Y, et al. Chem2Bio2RDF: a semantic framework for linking and data mining chemogenomic and systems chemical biology data. *BMC Bioinformatics* 2010;11:255 [FREE Full text] [doi: 10.1186/1471-2105-11-255] [Medline: 20478034]
24. Miñarro-Gimenez JA, Egaña Aranguren M, Martínez Béjar R, Fernández-Breis JT, Madrid M. Semantic integration of information about orthologs and diseases: the OGO system. *J Biomed Inform* 2011 Dec;44(6):1020-1031. [doi: 10.1016/j.jbi.2011.08.002] [Medline: 21864715]
25. Manola F, Miller E, RDF Core Working Group. World Wide Web Consortium. 2004 Feb 10. RDF Primer URL: <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/> [accessed 2012-02-05] [WebCite Cache ID 65E7C1ruR]
26. Corwin J, Silberschatz A, Miller PL, Marengo L. Dynamic tables: an architecture for managing evolving, heterogeneous biomedical data in relational database management systems. *J Am Med Inform Assoc* 2007 Feb;14(1):86-93 [FREE Full text] [doi: 10.1197/jamia.M2189] [Medline: 17068350]
27. Prud'hommeaux E, Seaborne A. World Wide Web Consortium. 2008 Jan 15. SPARQL Query Language for RDF URL: <http://www.w3.org/TR/rdf-sparql-query/> [accessed 2012-01-02] [WebCite Cache ID 64OiPASEa]
28. McGuinness DL, van Harmelen F, OWL Working Group. World Wide Web Consortium. 2004 Feb 10. OWL Web Ontology Language URL: <http://www.w3.org/TR/owl-features/> [accessed 2012-02-05] [WebCite Cache ID 65E7Hjm5l]
29. Schwaber K, Beedle M. *Agile Software Development With Scrum*. Upper Saddle River, NJ: Prentice Hall; 2002.
30. Millar MR. Tackling antibiotic resistance. International action required. *BMJ* 2010;340:c2978. [Medline: 20534657]
31. Piwowar HA, Becich MJ, Bilofsky H, Crowley RS, caBIG Data Sharing and Intellectual Capital Workspace. Towards a data sharing culture: recommendations for leadership from academic health centers. *PLoS Med* 2008 Sep 30;5(9):e183 [FREE Full text] [doi: 10.1371/journal.pmed.0050183] [Medline: 18767901]
32. Cruz IF, Xiao H. Ontology driven data integration in heterogeneous networks. *Stud Comput Intell* 2009(168):75-98. [doi: 10.1007/978-3-540-88075-2_4]
33. Schober D, Boeker M, Bullenkamp J, Huszka C, Depraetere K, Teodoro D, et al. The DebugIT core ontology: semantic integration of antibiotics resistance patterns. *Stud Health Technol Inform* 2010;160(Pt 2):1060-1064. [Medline: 20841846]
34. Teodoro D, Choquet R, Pasche E, Gobeill J, Daniel C, Ruch P, et al. Biomedical data management: a proposal framework. *Stud Health Technol Inform* 2009;150:175-179. [Medline: 19745292]
35. Teodoro D, Choquet R, Schober D, Mels G, Pasche E, Ruch P, et al. Interoperability driven integration of biomedical data sources. *Stud Health Technol Inform* 2011;169:185-189. [Medline: 21893739]
36. European Commission. 2010 Nov 4. A Comprehensive Approach on Personal Data Protection in the European Union URL: http://ec.europa.eu/justice/news/consulting_public/0006/com_2010_609_en.pdf [accessed 2012-01-03] [WebCite Cache ID 64Pi6bvNT]
37. Gamma E, Helm R, Johnson R, Vlissides J. *Design Patterns: Elements of Reusable Object-Oriented Software*. 1st edition. Reading, MA: Addison-Wesley; 1995.

38. Timpka T, Eriksson H, Gursky EA, Strömberg M, Holm E, Ekberg J, et al. Requirements and design of the PROSPER protocol for implementation of information infrastructures supporting pandemic response: a Nominal Group study. *PLoS One* 2011;6(3):e17941 [FREE Full text] [doi: [10.1371/journal.pone.0017941](https://doi.org/10.1371/journal.pone.0017941)] [Medline: [21464918](https://pubmed.ncbi.nlm.nih.gov/21464918/)]
39. Cheng CK, Ip DK, Cowling BJ, Ho LM, Leung GM, Lau EH. Digital dashboard design using multiple data streams for disease surveillance with influenza surveillance as an example. *J Med Internet Res* 2011;13(4):e85 [FREE Full text] [doi: [10.2196/jmir.1658](https://doi.org/10.2196/jmir.1658)] [Medline: [22001082](https://pubmed.ncbi.nlm.nih.gov/22001082/)]
40. Lung KR, Gorko MA, Llewelyn J, Wiggins N. Statistical method for the determination of equivalence of automated test procedures. *J Autom Methods Manag Chem* 2003;25(6):123-127. [doi: [10.1155/S146392460300021X](https://doi.org/10.1155/S146392460300021X)] [Medline: [18924621](https://pubmed.ncbi.nlm.nih.gov/18924621/)]
41. Johnston RJ, Duke JM. Benefit transfer equivalence tests with non-normal distributions. *Environ Resour Econ* 2007;41(1):1-23. [doi: [10.1007/s10640-007-9172-x](https://doi.org/10.1007/s10640-007-9172-x)]
42. Bizer C. D2R MAP-A DB to RDF Mapping Language. In: Proceedings. 2003 Presented at: 12th International World Wide Web Conference; May 24-24, 2003; Budapest, Hungary.
43. Ruch P, Gobeill J, Lovis C, Geissbühler A. Automatic medical encoding with SNOMED categories. *BMC Med Inform Decis Mak* 2008;8 Suppl 1:S6 [FREE Full text] [doi: [10.1186/1472-6947-8-S1-S6](https://doi.org/10.1186/1472-6947-8-S1-S6)] [Medline: [19007443](https://pubmed.ncbi.nlm.nih.gov/19007443/)]
44. Ruch P. Automatic assignment of biomedical categories: toward a generic approach. *Bioinformatics* 2006 Mar 15;22(6):658-664 [FREE Full text] [doi: [10.1093/bioinformatics/bti783](https://doi.org/10.1093/bioinformatics/bti783)] [Medline: [16287934](https://pubmed.ncbi.nlm.nih.gov/16287934/)]
45. W3C Semantic Web Deployment Working Group. World Wide Web Consortium. 2012. SKOS Simple Knowledge Organization System URL: <http://www.w3.org/2004/02/skos/> [accessed 2012-02-05] [WebCite Cache ID 65E7M0Ek2]
46. Goble C, Stevens R. State of the nation in data integration for bioinformatics. *J Biomed Inform* 2008 Oct;41(5):687-693. [doi: [10.1016/j.jbi.2008.01.008](https://doi.org/10.1016/j.jbi.2008.01.008)] [Medline: [18358788](https://pubmed.ncbi.nlm.nih.gov/18358788/)]
47. Momtchev V, Peychev D, Primov T, Georgiev G. International Semantic Web Challenge. 2009. Expanding the Pathway and Interaction Knowledge in Linked Life Data URL: <http://www.cs.vu.nl/~pmika/swc/documents/Linked%2520Life%2520Data-LLD%2520semantic%2520web%2520challenge%25202009.pdf> [accessed 2012-05-09] [WebCite Cache ID 67X4tHYmR]
48. Morsey M, Lehmann J, Auer S, Ngomo ACN. DBpedia SPARQL benchmark: performance assessment with real queries on real data. In: Proceedings. 2011 Presented at: 10th International Semantic Web Conference; Oct 23-27, 2011; Bonn, Germany p. 454-469.
49. Schmidt M, Hornung T, Lausen G, Pinkel C. SP²Bench: a SPARQL performance benchmark. In: Proceedings. 2009 Presented at: IEEE 25th International Conference on Data Engineering (ICDE '09); Mar 29-Apr 2, 2009; Shanghai, China. [doi: [10.1109/ICDE.2009.28](https://doi.org/10.1109/ICDE.2009.28)]
50. Lee F, Teich JM, Spurr CD, Bates DW. Implementation of physician order entry: user satisfaction and self-reported usage patterns. *J Am Med Inform Assoc* 1996 Feb;3(1):42-55 [FREE Full text] [Medline: [8750389](https://pubmed.ncbi.nlm.nih.gov/8750389/)]
51. Reynolds R, Hope R, Williams L, BSAC Working Parties on Resistance Surveillance. Survey, laboratory and statistical methods for the BSAC Resistance Surveillance Programmes. *J Antimicrob Chemother* 2008 Nov;62 Suppl 2:ii15-ii28 [FREE Full text] [doi: [10.1093/jac/dkn349](https://doi.org/10.1093/jac/dkn349)] [Medline: [18819976](https://pubmed.ncbi.nlm.nih.gov/18819976/)]
52. Calvanese D, Giacomo G, Lenzerini M. Ontology of integration and integration of ontologies. In: Proceedings. 2001 Presented at: 2001 International Workshop on Description Logic; Aug 1-3, 2001; Stanford, CA, USA.

Abbreviations

ARTEMIS: Antimicrobial Resistance Trend Monitoring System

CI: confidence interval

DebugIT: Detecting and Eliminating Bacteria Using Information Technology

EARS-Net: European Antimicrobial Resistance Surveillance Network

ICDR: local clinical data repository

RDF: Resource Description Framework

SEARCH: Sentinel Surveillance of Antibiotic Resistance in Switzerland

SKOS: Simple Knowledge Organization System

SNOMED-CT: Systematized Nomenclature of Medicine-Clinical Terms

SPARQL: Simple Protocol and RDF Query Language

WHO: World Health Organization

WHO-ATC: World Health Organization- Anatomical Therapeutic Chemical

Edited by G Eysenbach; submitted 07.01.12; peer-reviewed by L Balkanyi, T Timpka; comments to author 29.01.12; revised version received 05.03.12; accepted 29.04.12; published 29.05.12.

Please cite as:

Teodoro D, Pasche E, Gobeill J, Emonet S, Ruch P, Lovis C

Building a Transnational Biosurveillance Network Using Semantic Web Technologies: Requirements, Design, and Preliminary Evaluation

J Med Internet Res 2012;14(3):e73

URL: <http://www.jmir.org/2012/3/e73/>

doi: [10.2196/jmir.2043](https://doi.org/10.2196/jmir.2043)

PMID: [22642960](https://pubmed.ncbi.nlm.nih.gov/22642960/)

©Douglas Teodoro, Emilie Pasche, Julien Gobeill, Stéphane Emonet, Patrick Ruch, Christian Lovis. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.05.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Using Internet Search Engines to Obtain Medical Information: A Comparative Study

Liupu Wang^{1,2,3}, PhD; Juexin Wang¹, MS; Michael Wang⁴, MD; Yong Li², PhD; Yanchun Liang¹, PhD; Dong Xu⁵, PhD

¹Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun, China

²School of Mathematics, Jilin University, Changchun, China

³Department of Computer Science and Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, United States

⁴Department of Pathology and Anatomical Sciences and Ellis Fischel Cancer Center, University of Missouri, Columbia, MO, United States

⁵Department of Computer Science, University of Missouri, Columbia, MO, United States

Corresponding Author:

Dong Xu, PhD

Department of Computer Science

University of Missouri

201 Engineering Building West

Columbia, MO,

United States

Phone: 1 573 884 1887

Fax: 1 573 882 8318

Email: xudong@missouri.edu

Abstract

Background: The Internet has become one of the most important means to obtain health and medical information. It is often the first step in checking for basic information about a disease and its treatment. The search results are often useful to general users. Various search engines such as Google, Yahoo!, Bing, and Ask.com can play an important role in obtaining medical information for both medical professionals and lay people. However, the usability and effectiveness of various search engines for medical information have not been comprehensively compared and evaluated.

Objective: To compare major Internet search engines in their usability of obtaining medical and health information.

Methods: We applied usability testing as a software engineering technique and a standard industry practice to compare the four major search engines (Google, Yahoo!, Bing, and Ask.com) in obtaining health and medical information. For this purpose, we searched the keyword *breast cancer* in Google, Yahoo!, Bing, and Ask.com and saved the results of the top 200 links from each search engine. We combined nonredundant links from the four search engines and gave them to volunteer users in an alphabetical order. The volunteer users evaluated the websites and scored each website from 0 to 10 (lowest to highest) based on the usefulness of the content relevant to breast cancer. A medical expert identified six well-known websites related to breast cancer in advance as standards. We also used five keywords associated with breast cancer defined in the latest release of Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) and analyzed their occurrence in the websites.

Results: Each search engine provided rich information related to breast cancer in the search results. All six standard websites were among the top 30 in search results of all four search engines. Google had the best search validity (in terms of whether a website could be opened), followed by Bing, Ask.com, and Yahoo!. The search results highly overlapped between the search engines, and the overlap between any two search engines was about half or more. On the other hand, each search engine emphasized various types of content differently. In terms of user satisfaction analysis, volunteer users scored Bing the highest for its usefulness, followed by Yahoo!, Google, and Ask.com.

Conclusions: Google, Yahoo!, Bing, and Ask.com are by and large effective search engines for helping lay users get health and medical information. Nevertheless, the current ranking methods have some pitfalls and there is room for improvement to help users get more accurate and useful information. We suggest that search engine users explore multiple search engines to search different types of health information and medical knowledge for their own needs and get a professional consultation if necessary.

KEYWORDS

Internet search; page rank; Google, Yahoo!, Bing, Ask.com; medical information; health information seeking; breast cancer; SNOMED CT; user experience evaluation; usability testing; hallway testing; software engineering

Introduction

The Internet is becoming one of the most important sources to obtain medical and health information for the general public [1]. The Pew Internet & American Life Project reported [2] that about 80% of Internet users look for medical or health-related information through the Internet. It has been found that the most common health and medical topics searched on the Internet were a specific disease or medical problem, certain medical treatment, diet, nutrition and vitamins, and exercise or fitness [2], as well as prescription drugs [3]. To help the public obtain accurate and useful medical information, the US National Library of Medicine developed an official website (MedlinePlus; www.nlm.nih.gov/medlineplus/) to provide consumers with access to current full-text publications in health and medicine [4]. Volunteers from universities, hospitals, and research institutions summarize their work in the health sections of Wikipedia (www.wikipedia.org). The editors have invited the medical community to edit and update the content in Wikipedia to provide reliable and understandable health information [5]. It was suggested that such a World Wide Web-based supporting system would become an important professional tool for future medicine [6]. Researchers also explored the feasibility of using online medical information as a diagnostic aid [7]. An online patient recruitment company studied patients' search queries and questions for clinical trial information in order to improve the clinical trial recruitment process [8]. Thus, the Internet has become an indispensable source for the public, patients, and health care professionals to obtain information about health, diseases, and medical treatment.

Various investigations have been conducted for improving search methods to help users obtain accurate and useful information [9]. Some studies have improved the search results or efficiency by analyzing the purpose of the Internet search [10]. In particular, researchers have classified and analyzed the information needs of users to improve search engines [11]. A medicine-related information search is different from other information searches, since users often use medical terminology, disease knowledge, treatment options, and so on. Hence, some studies targeted the characteristics and motivations of health information seekers and the usage of online medical information [12,13]. As early as 1997, Bonati's group studied the reliability of health care information from two major search engines, Yahoo! and Excite. They found that only a few websites provided the comprehensive and accurate information sought due to lack of mature techniques of most search engines used for public health care information [14]. In 1999, researchers compared the accuracy and reliability of search results from five search engines: AltaVista, Excite, Hotbot, Infoseek, and Lycos [15]. A similar study compared two major search engines, Alltheweb and Excite [16]. Their results demonstrated that the use of the general search engines for health care information,

as well as the use of specialized medical and health websites, had dramatically increased. Another area of study is the user experience in health care information search [17]. In 2005, Forrester Research, Inc. conducted a Web survey to assess search engines by evaluating AOL, Google, MSN, and Yahoo! according to 11 parameters in user experience [18]. However, it is difficult to obtain a comprehensive evaluation for the search engines based only on these scores without a medical professional's input.

While several major search engines are available, most users limit their Internet search to one search engine. This search habit raises several questions regarding the search for medical information. Does a single search engine provide reliable medical information? Are search results from different engines similar when using multiple search engines? Does searching multiple search engines for the same queries add value? Is any search engine significantly better than the others? These questions become more relevant as the market shares of Internet search engines are evolving. A report from comScore, a global company measuring the digital world and the preferred source of digital marketing intelligence, showed that in September 2009, the top five search engines on the market were Google, Yahoo!, Microsoft Sites, Ask Network, and AOL LLC [19]. According to comScore, Google has about two-thirds of the US market share in Internet searching, and its share is increasing. Then a natural question is whether Google is enough for medical information searching without needing to check other search engines.

Among all the medical search queries on the Internet, cancer-related information is one of the most popular topics. In particular, breast cancer, as the most common cancer and the second-leading cause of cancer deaths among American women, draws much of the public's attention, especially on the Internet [20]. In a previous study [21], we recruited some Internet users and a medical professional, and asked them to score the content of webpages resulting from Google searches to reflect their user experience quantitatively from a software engineering perspective. This pilot study showed that usability testing is very useful for evaluating Internet searches in obtaining medical information. It also showed that the specificity of Google in searching for medical information is often satisfactory. However, the study targeted only one search engine, Google. The current study extended our early work by adding Yahoo!, Bing, and Ask.com, which allowed us to conduct a comparative study. Among the four search engines, Bing is a new one built in 2009 by Microsoft. We still chose *breast cancer* as a keyword and Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) [22,23] as a standard reference in our validation process. We used Java Vector data structure to read and record the ranking, website name, and URL of webpages in the search result list.

Methods

Design

We used the following protocol to conduct our research. First, search *breast cancer* in Google, Yahoo!, Bing, and Ask.com and save the top 200 websites shown in each search, including website name, URL, and ranking of each result in every group. Second, apply the Google PageRank tool to check and record the Google PageRank for each search result in every group. Third, compare website links to check overlapping of search results from the same search engine and between the different search engines. Fourth, define standards of the most helpful and commonly used websites for breast cancer by a breast cancer research expert, and have eight volunteers study the standard websites so that they would know what kinds of websites are informative. Fifth, ask the eight volunteers to score each of the combined, nonredundant websites in alphabetical order, scoring from 0 (a useless website) to 10 (the most helpful website). We would then collect the data from volunteers and rerank the websites from high to low in every group (corresponding to each search engine). Sixth, classify each search result according to its webpage content, with four types of content defined. Seventh, use Java.net to retrieve the text of all nonredundant websites, and use the subpages to study the pattern of usage of the keyword *breast cancer* and its related SNOMED CT terms. Eighth, analyze the data and compare the effectiveness of Google, Yahoo!, Bing, and Ask.com in obtaining medical and health information.

Collection of Search Results

On October 27, 2009, we used four search engines—Google, Yahoo!, Bing, and Ask.com—to search *breast cancer* with the default search parameters. A Java application programming interface named JExcelApi [24] was used to record related data

into an Excel (Microsoft Corporation, Redmond, WA, USA) file automatically. We obtained the search results using the Windows 7 Professional Service Pack 1 (Microsoft Corporation) operating system in English. All the searches and their evaluations were carried out on machines physically located in Columbia, Missouri, USA.

We obtained a total of 798 search results in the following four groups. Google (www.google.com) obtained 40,800,000 search results, and the response time was 0.13 seconds. We chose the top 200 websites from these results to compose the Google sample group in this study. Yahoo! (www.search.yahoo.com) obtained 262,000,000 search results. We chose the top 200 websites from these results as the Yahoo! sample group. Bing (www.bing.com) obtained 74,500,000 search results. We chose the top 200 websites from these results as the Bing sample group. Ask.com (www.ask.com) had 9,080,000 search results. We chose all 198 websites available to users for the Ask.com sample group. Ask.com is a metasearch engine, which aggregates and selects the results from several other search engines into a single list.

Then we combined the nonredundant links and provided them to volunteer users in alphabetical order, as Figure 1 shows. Redundancy is defined by websites having the same URL address. Sometimes, different URLs or even domain names have identical or similar content. However, such cases are relatively rare and it is very time consuming to manually label them. Hence, we did not consider this issue in this study. We used alphabetical order so as to avoid the bias from ranked order, as volunteers may unconsciously have followed the search engines' ranking of the search results. Using nonredundant links reduced the volunteers' workload, from 798 search results to 592 Web links. At the same time, we also grouped the sample search results together if they share the same domain name for the user's convenience.

Figure 1. Sample of score list used by volunteers.

Score	Web Url
	p://aabcainc.org
0	p://aarp.org/.../harvard_breast-cancer-introduction_1.html
1	p://abcnews.go.com/Health/OnCallPlus
2	p://abcnews.go.com/Health/OnCallPlus/story?id=3666505&page=1
3	p://abcnews.go.com/Health/OnCallPlusBreastCancerNews/
4	
5	p://adam.about.com/encyclopedia/infectiousdiseases/Breast-cancer.htm
6	http://adam.about.com/reports/Breast-cancer.htm
7	
	http://answers.yahoo.com/question/index?qid=20061031181945AAPNACz
	http://answers.yahoo.com/question/index?qid=20070708173453AA9fRLM
	http://apps.nccd.cdc.gov/cancercontacts/nbccedp/contacts.asp
	http://bbc.co.uk/health/womens_health/issues_breastcancer2.shtml
	http://bcresourcedirectory.org
	http://bcsupport.org
	http://blog.newsweek.com/blogs/thehumancondition/archive/2009/10/02/sexy-breast-cancer-ads-provocative-or-patronizing-.aspx
	http://breastcancer.about.com
	http://breastcancer.about.com/od/definition/a/bc_definition.htm
	http://breastcancer.about.com/od/whatisbreastcancer/p/symptoms.htm
	http://breastcancer.bellaonline.com
	http://breastcancer.care2.com
	http://breastcancer.com
	http://breastcancer.einnews.com
	http://breast-cancer.emedtv.com
	http://breastcancer.uams.edu
	http://breastcanceranalysis.com
	http://breastcancerangels.org
	http://breastcancercured.com

PageRank Value of Search Results

The Google search engine has a prominent status and a large market share in Internet search. The PageRank algorithm is one of Google's core technologies. The PageRank value is a parameter to evaluate the importance of a website, ranging from 0 to 10 [25]. Many software tools for research or business use the Google PageRank technology. We used the PageRank Checker Tool [26] and the Google PageRank Checker [27] to evaluate and record the Google PageRank value for each search result in every group.

Scoring Criteria

The volunteers evaluated each website from 0 to 10 based on its usefulness concerning breast cancer, where a rating of 10 indicated the most helpful website. To orientate the volunteers, Dr Michael Wang, a physician and a breast cancer researcher, defined gold standard websites that provide the most useful information about breast cancer to the public. These websites usually have all the basic information about breast cancer, such as understanding breast cancer, symptoms, and diagnosis. It is convenient for the public to obtain the information they need rapidly from these websites. He identified six gold standard websites: standard No. 1, from the US National Cancer Institute; standard No. 2, from the American Cancer Society; standard No. 3, from the Mayo Clinic; standard No. 4, from MedicineNet; standard No. 5, Wikipedia; and standard No. 6, from Susan G. Komen for the Cure.

Evaluation of Websites by Volunteers

We provided a list of nonredundant search results to eight nonphysician volunteers. Each volunteer had at least 5 years' experience of using the Internet. The volunteers first studied the six standard websites so that they knew what they could

expect from high-quality websites, annotated by the expert. However, they did not have to give these sites high scores if they did not regard them as helpful. Then they read the sample websites individually and scored them based on the standard and their own experience. The eight volunteers in this study were from the University of Missouri-Columbia: volunteer 1 was a 30-year-old female PhD student with biochemistry background; volunteer 2 was a 36-year-old male Associate Professor with a PhD in computer science; volunteer 3 was a 31-year-old male PhD student in computer science; volunteer 4 was a 29-year-old male PhD student in computer science; volunteer 5 was a 23-year-old female undergraduate student in biochemistry; volunteer 6 was a 27-year-old male graduate student in chemistry; volunteer 7 was a 25-year-old male graduate student in economics; and volunteer 8 was a 22-year-old female undergraduate student in biological science.

These volunteers were frequent Internet users but did not have any background in medical science. The method of choosing and inviting testing users is different from the regular way of randomly choosing volunteers in hallway testing [28]. A typical hallway testing limits the workload of each volunteer to 10 minutes with a questionnaire containing no more than 50 questions. In this study each volunteer evaluated 592 sample websites. Browsing and scoring all of them took about 10 hours per volunteer on average. Given the heavy workload, we made sure that each volunteer was committed to participating in the study. During the study, we advised each volunteer to do the study whenever he or she could find some time but to finish the evaluation as soon as possible to reduce the negative effect caused by the expiration of some Web links. It took from 3 to 17 days to finish evaluating all the webpages of the search results. Since many of the search results contained medical terminologies and related knowledge, we only selected

volunteers with a higher-education background so that they could understand the content. The volunteers' experience in Internet use ensured that they could conduct the study smoothly [29]. The users' age and gender were diverse, following early suggestions [30,31]. Overall we attempted a representative study given the availability of volunteers.

Categorization of Search Results

We studied the search results by classifying their webpage content, as different categories of content may have an impact on users' scores. We manually classified 798 search results into the following four types: type 1, websites for basic breast cancer knowledge targeting the general population such as news sites, Web tribunes, personal websites, and blogs; type 2, nonprofit organization websites for breast cancer patients and their families including websites of breast cancer societies and foundations; type 3, corporate websites for consumers including advertisement websites for medicine, devices, products, and so on; and type 4, websites for breast cancer professionals and researchers such as websites from universities, research institutions, hospitals, and government.

Ranking Based on Selected Keywords

We provided a reference ranking based on the occurrence of selected keywords. We used Java.net to obtain the text content automatically from all websites. All the text in a website in the search result formed the main page text corpus. All the subpages of this website combined with corpus A formed the main and subpages text corpus. Based on the description of breast cancer in SNOMED CT, we selected four more keywords: *malignant tumor*, *neoplasm*, *sarcoma*, and *carcinoma*. We reranked the websites in each group based on the frequency of these keywords in the text corpus according to the following rules: (1) the website with the most types of keywords is ranked on top, (2) if 1 is inconclusive, the one with greater occurrence of *breast cancer* is on top, (3) if 2 is inconclusive, the one with

more hits of the other three key words is on top, and (4) if 3 is inconclusive, the one with higher ranking in the original search result list is on top.

Results

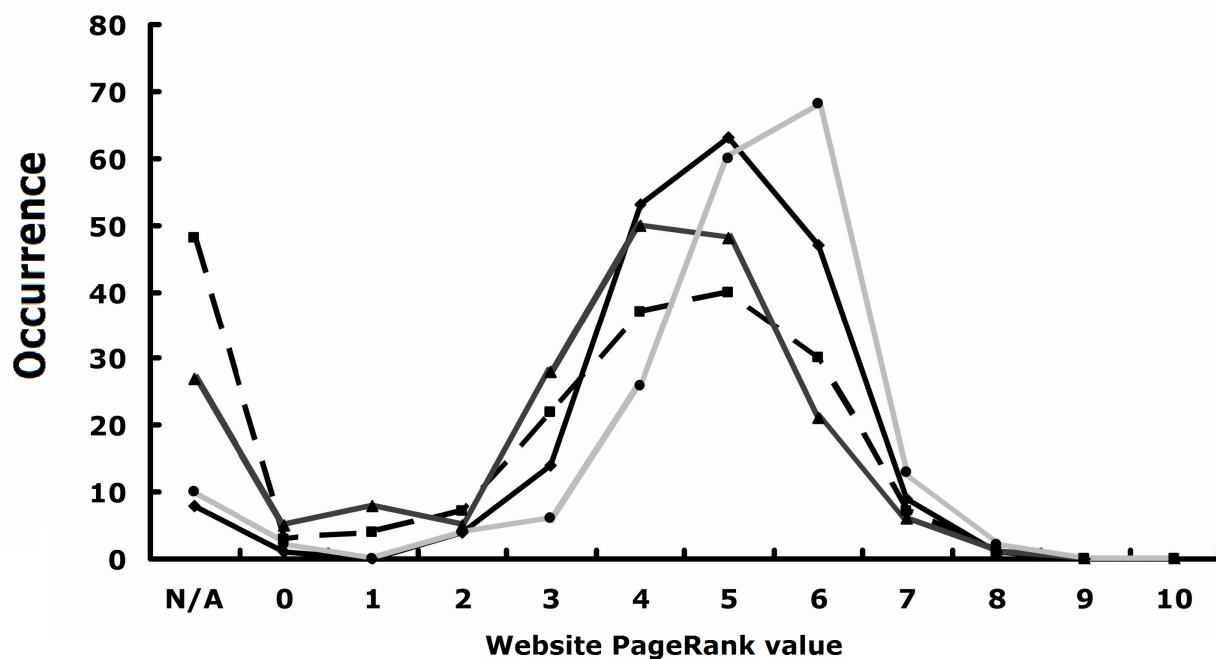
Validity of Search Results

Search validity is defined according to whether a user can successfully open the URL of a search result. Our analysis showed that the validities of the search results from the four search engines were significantly different. The validity of Google search results was 100% (200/200), while Yahoo! had a validity of 92.5% (185/200) with 15 invalid results (websites 21, 53, 75, 85, 118, 119, 120, 126, 129, 140, 147, 149, 162, 171, and 176 in [Multimedia Appendix 1, Table 1](#)), Bing had a validity of 98.5% (197/200) with three invalid results (74, 148, and 149), and Ask.com had a validity of 98.9% (196/198) with two invalid results (12 and 83). This indicates that Google probably updates its database of search results more frequently than other search engines. We did not use the cached pages of the invalid sites for any further study. The cached pages are typically incomplete and they are often ignored by users. Hence, they are not suitable for usability studies.

PageRank Value of Search Results

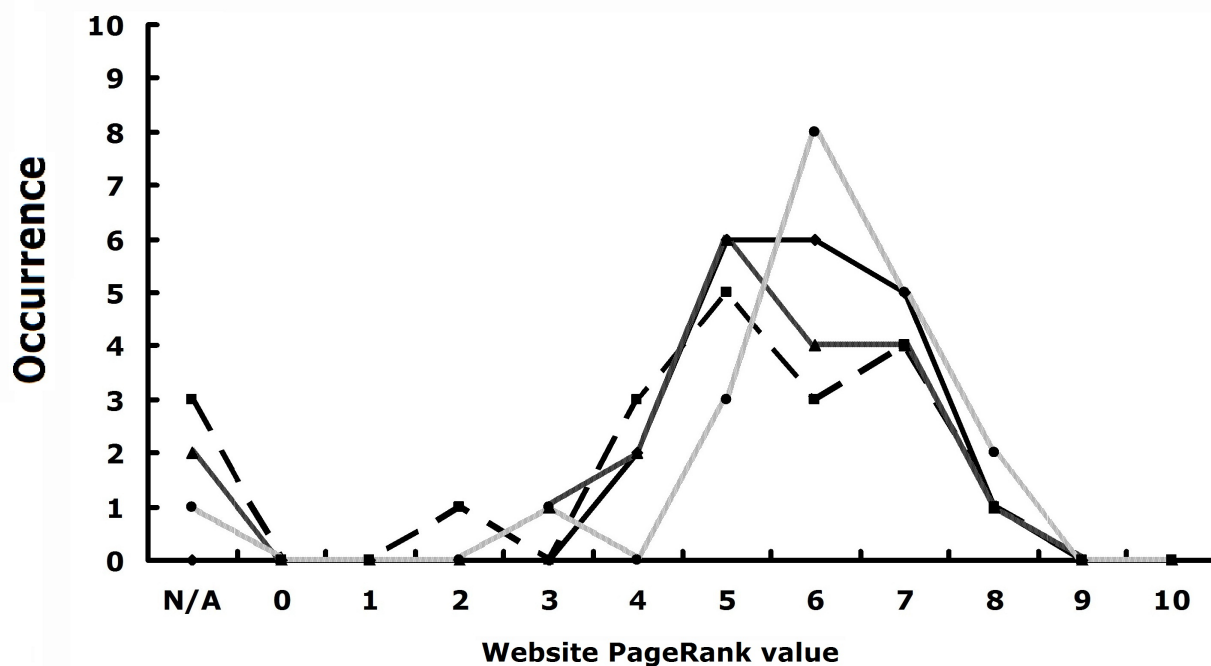
[Figure 2a](#) shows the PageRank distribution of search results in each of the four groups. The distributions of the four groups were similar, although Ask.com skewed more toward higher PageRank values than the other three. It is likely that Ask.com as a metasearch engine selects the results from other search engines by considering a factor similar to PageRank. Since many Internet search users look at only the top 10 or 20 hits, we also compared the search engines using the top 20 search results. [Figure 2b](#). It shows that the features are almost the same as those in the top 200 hits.

Figure 2. Distribution of PageRank values of search results from Google, Yahoo!, Bing, and Ask.com. (a) All search results, (b) top 20 search results. N/A = not available.



—●— Google —■— Yahoo! —▲— Bing —◆— Ask.com

(a)



—●— Google —■— Yahoo! —▲— Bing —◆— Ask.com

(b)

Ranking of Six Standard Websites

For the six standard websites, four search engine hits contained the exact URLs for five of them. The other one (standard No. 2, <http://www.cancer.org/docroot/home/index.asp>) was identified by two search engines (Bing and Ask.com), and its subpage

(http://www.cancer.org/docroot/cricri_2x.asp?sitearea=lm&dt=5) appeared in the Google and Yahoo! search results as Google No. 16 and Yahoo! No. 1, respectively. Counting this subpage, all six standard websites appeared on the top 30 of each search engine's results, except for one (standard No. 3 ranked 33 by Bing), as Table 1 shows.

Table 1. Ranking of six standard websites.

Web-site No.	Website name	Website URL	Search engine ranking			
			Google	Yahoo!	Bing	Ask.com
1	National Cancer Institute: Breast cancer	http://www.cancer.gov/cancer-topics/types/breast (not archived)	4	9	3	7
2	American Cancer Society: Information and Resources for Cancer	http://www.cancer.org/docroot/home/index.asp (not archived)	16	1	5	13
3	MayoClinic: Breast cancer	http://www.mayoclinic.com/health/breast-cancer/DS00328 (http://www.webcitation.org/67hgmdCNF)	9	17	33	22
4	MedicineNet.com: Breast cancer	http://www.medicinenet.com/breast_cancer/article.htm (http://www.webcitation.org/67hgtCTA8)	8	26	4	9
5	Wikipedia: Breast cancer	http://en.wikipedia.org/wiki/Breast_cancer (http://www.webcitation.org/67hh3W7P8)	13	23	2	5
6	Susan G. Komen Breast Cancer Foundation (http://www.webcitation.org/67hh7GieT)	http://ww5.komen.org/	3	10	8	6

Overlap Between Search Results

We found that only 397 of 798 (49.8%) search results were nonredundant, with 401 results having a duplicate URL. In addition, 466 of 798 (58.4%) sample results had redundant domain names. As Table 2 shows, within each group Google search results had no redundant URLs, but 18 of the 200

websites (9%) had domain names redundant with other search results in the same group. This is much less than in the other three groups, where Yahoo!, Ask.com, and Bing had domain redundancy of 12% (24/200), 16% (31/198), and 16% (32/200), respectively. Too much redundancy may lead to poor user experience. Thus, it is an important factor in user experience evaluation.

Table 2. Redundancy of search results.

Search engine	Google (n = 200)		Yahoo! (n = 200)		Bing (n = 200)		Ask.com (n = 198)	
	URL redundancy	Domain redundancy	URL redundancy	Domain redundancy	URL redundancy	Domain redundancy	URL redundancy	Domain redundancy
Google	0	18	61	75	49	67	67	79
Yahoo!	61	75	2	24	60	76	42	52
Bing	49	67	60	76	2	32	38	61
Ask.com	67	79	42	52	38	61	13	31

In Table 2, as an example, Google and Yahoo! shared 61 results with the same URLs and 75 results with the same domain names.

Categorization Distribution of Search Results

Figure 3 and Figure 4 show that the four search engines differed in their distributions of search results in terms of content types. To review, the four types of websites are (1) basic knowledge websites targeting the general population, (2) nonprofit organization websites, (3) corporation websites, and (4) websites for professionals and researchers. Overall, most results fell into

type 1, type 2, and type 4, with proportions all around 30%. The distributions of types 1–4 were 29.7% (237/798), 29.2% (233/798), 12% (93/798), and 29.4% (235/798), respectively. Google mostly covered type 1 and type 2, with 33% (66/200) and 34% (67/200) of all Google search results, respectively. Yahoo! had overrepresentation in type 1 (33% (66/200) among all Yahoo! search results). Bing emphasized type 1 and type 4 (35% (70/200) and 37% (73/200) among all Bing search results, respectively). Ask.com focused on type 2 (40% (79/198) among all Ask.com search results).

Figure 3. Distributions of the four webpage types among all search results and in the four search engines.

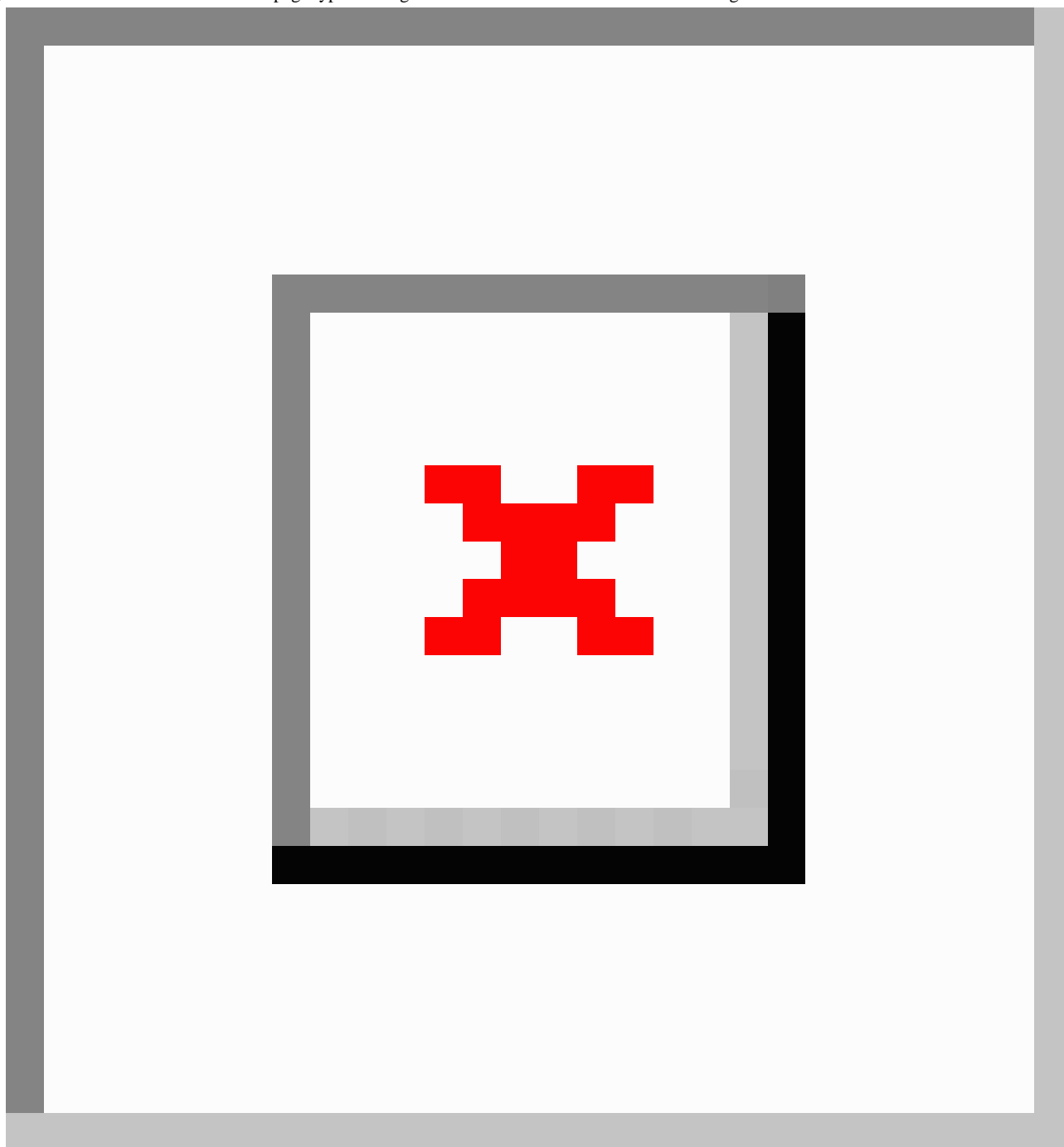
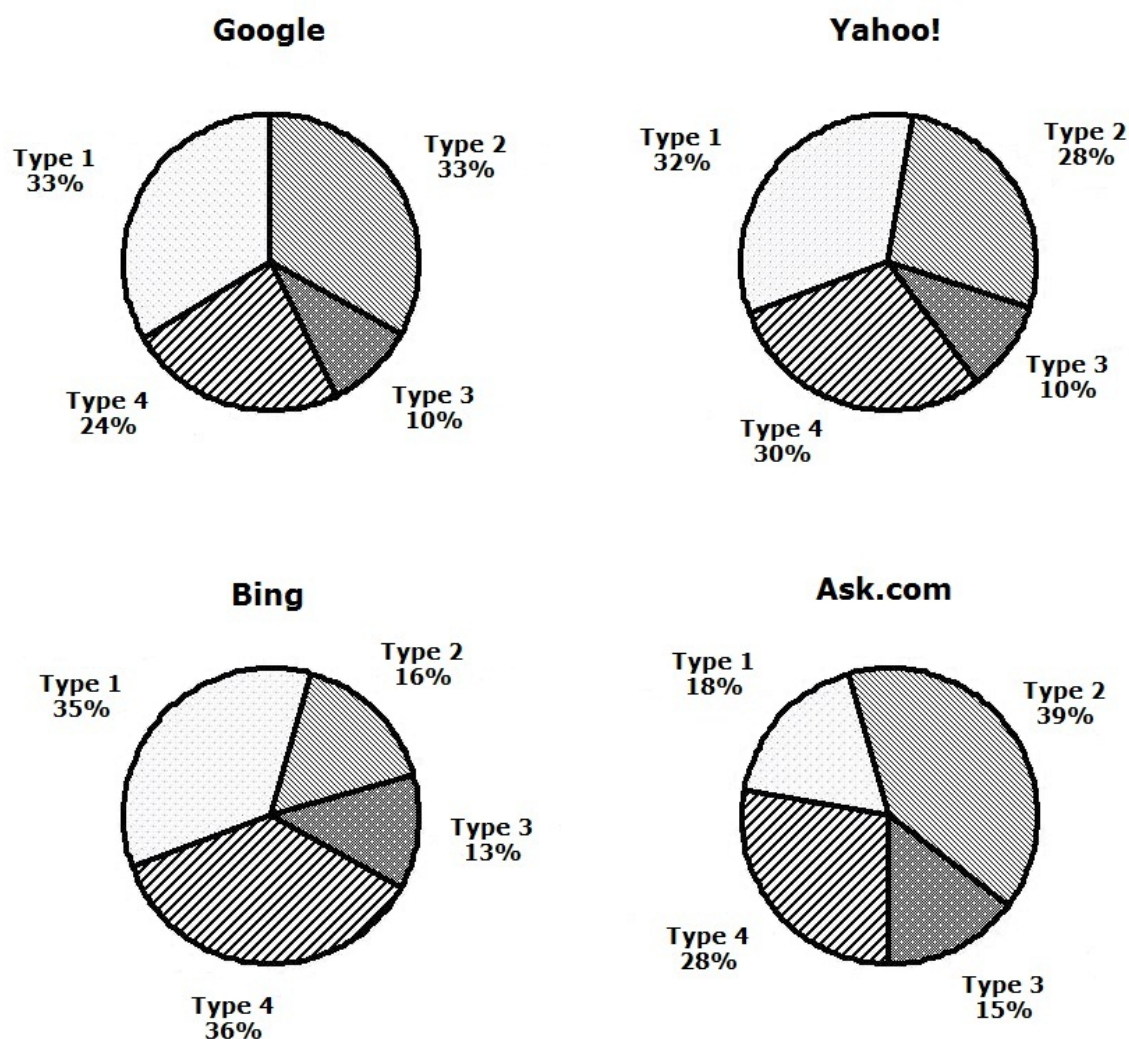
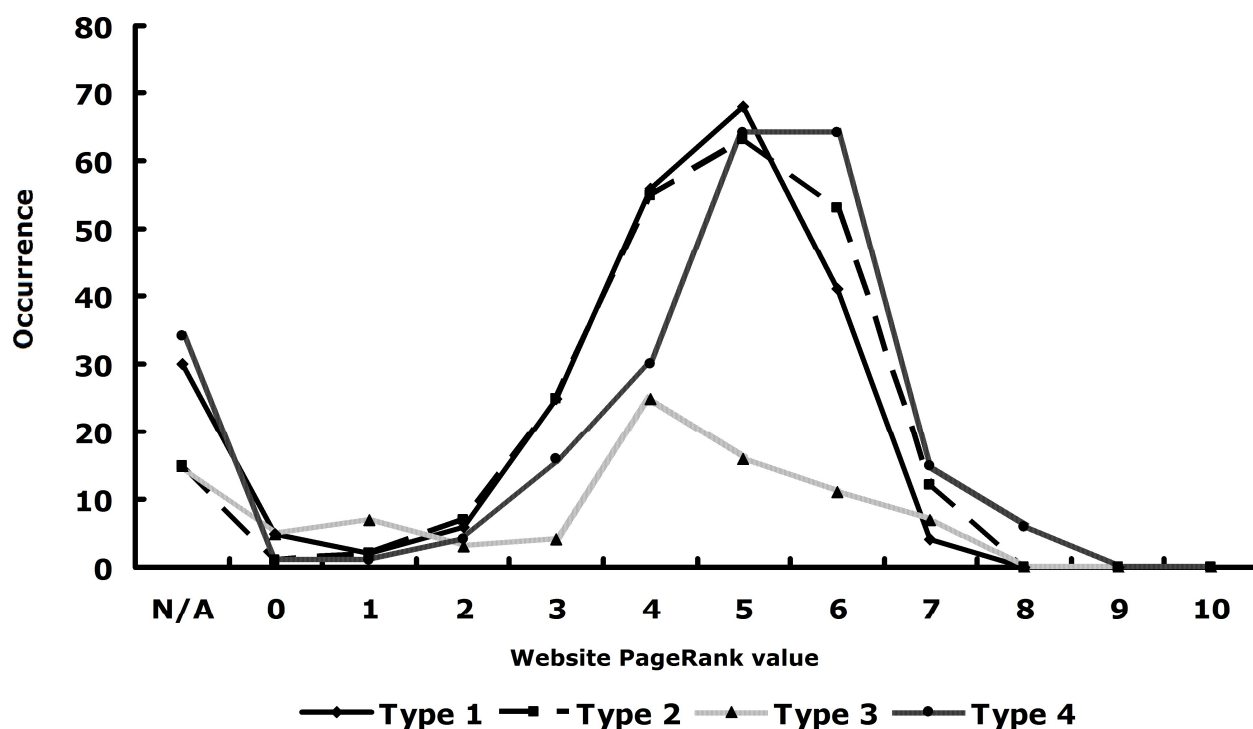


Figure 4. Distribution of the four webpage types in each search results group (by search engine).

PageRank Values of Four Webpage Types

Figure 5 shows the distribution of the PageRank value among the different types of websites. The average PageRank values of the four types are 4.48 for type 1, 4.72 for type 2, 4.03 for type 3, and 5.14 for type 4. It is understandable that type 4 has the highest PageRank value, as these websites contain authoritative content from breast cancer professionals and researchers. Because there are more personal websites (instead

of institutional websites) in type 1, the popularity of these websites tends to be lower, and hence type-1 websites overall have lower PageRank values. Since type 3 (corporate websites) is mainly related to advertisement, users might tend to avoid these websites, and hence they also have lower PageRank values overall. Many of the type-2 (nonprofit) websites with high PageRank values belong to the same group—for example, the American charity website CharityUSA.com.

Figure 5. PageRank value of the four webpage types. N/A = not available.

Volunteers' Scores for Standard Websites

While the volunteers were given the standard websites, they did not have to give the sites a score of 10. While some sites had high scores (eg, the volunteers gave standard No. 5 an average score of 9.88), it is interesting that volunteers often

gave some sites relatively low scores, as Table 3 shows. In particular, standard No. 2 (<http://www.cancer.org/docroot/home/index.asp>) had an average score of only 6.50; this website was the only one that was not identified by all 4 search engines. This shows the due diligence of the volunteers.

Table 3. Volunteers' scores for the six standard websites.

Website No.	Volunteer No.								Average score
	1	2	3	4	5	6	7	8	
1	10	9	10	9	8	8	10	6	8.75
2	6	8	8	6	5	7	7	5	6.50
3	10	8	10	9	9	9	10	5	8.75
4	10	10	10	9	9	8	9	8	9.13
5	10	10	10	10	10	10	10	9	9.88
6	10	9	10	8	10	7	10	6	8.75

Scores of Search Results From Different Search Engines

Figure 6a shows the score distribution in each search group. Distribution patterns were similar in Google and Bing, and in Yahoo! and Ask.com. Bing had the highest average score at 5.70, with more high scores from 7 to 10, while Yahoo!, Google, and Ask.com had average scores of 5.07, 4.78, and 4.10, respectively. Figure 6b shows the score distribution for top 20 search results in each search group. Bing again had the highest average score at 7.14, while Google, Ask.com, and Yahoo! had average scores of 6.85, 6.36, and 5.77, respectively.

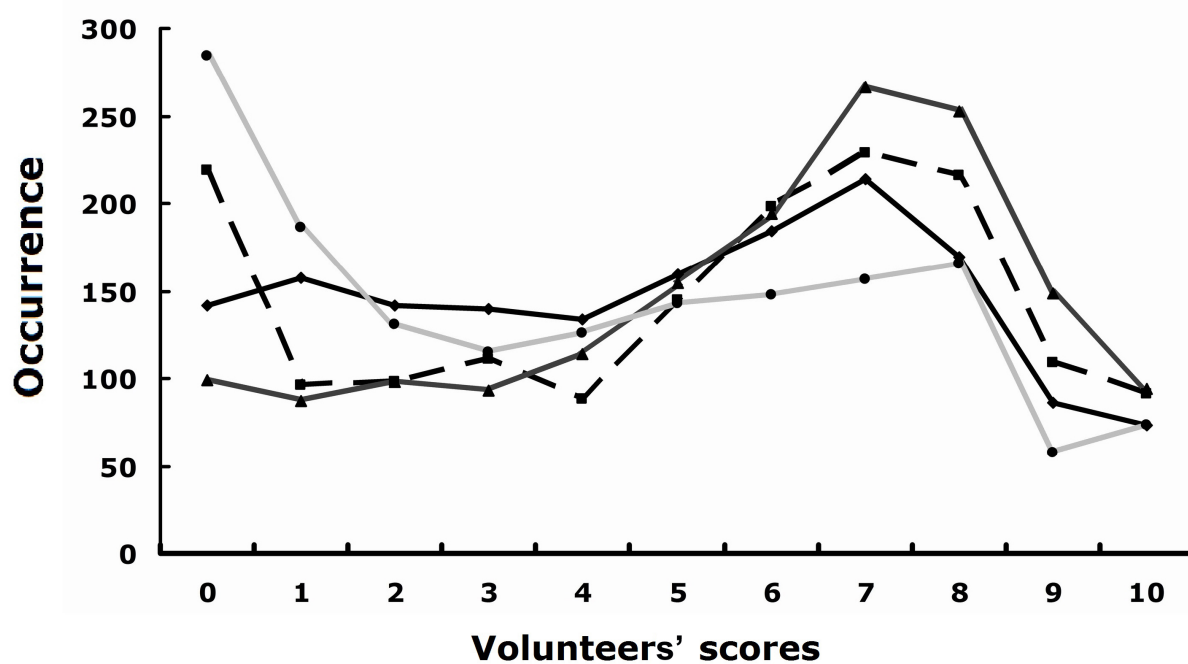
Figure 7 shows the average score versus ranking with a window size of 20. Table 4 provides some additional assessment. The general trend of the curves in Figure 7 is as expected—that is, the higher the ranking, the higher the volunteer's score. In terms of the correlation between the scores and the ranking by the search engine, the best search engine was Ask.com, followed by Google, Bing, and Yahoo!. A minor issue in this comparison is that each search engine had some redundant search results or domain names. Given that the redundancy was not very high, we ignored the issue for the comparison between different search engines.

Table 4. Performance (average score) of the four search groups.

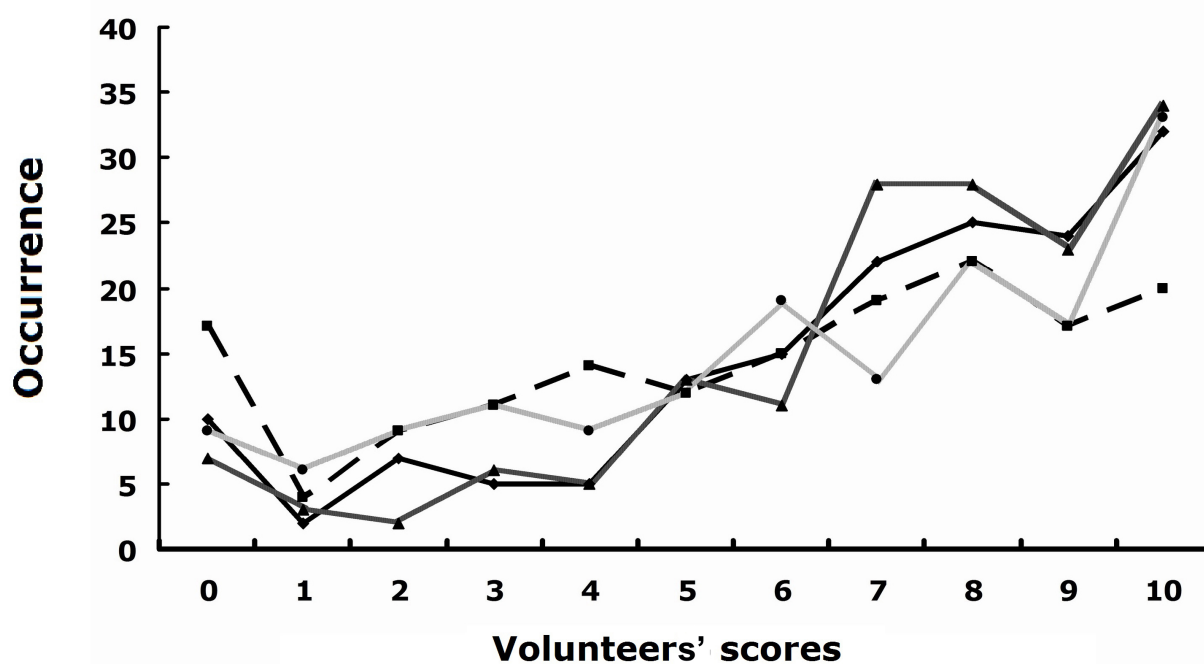
Performance measure	Google	Yahoo!	Bing	Ask.com
Top 10 websites	7.34	5.46	7.70	6.79
Top 20 websites	6.85	5.77	7.14	6.36
Top 50 websites	5.96	5.85	6.19	5.65
Top 100 websites	5.28	5.49	6.08	5.19
Total	4.78	5.07	5.70	4.14
Pearson correlation coefficient ^a	-.3036	-.1937	-.1964	-.5099
Spearman rank correlation ^a	-.3062	-.2051	-.2281	-.4725

^a Pearson correlation coefficient and Spearman rank correlation coefficient are between volunteer's score and ranking by the search engine. The larger the absolute value, the greater the correlation in this case.

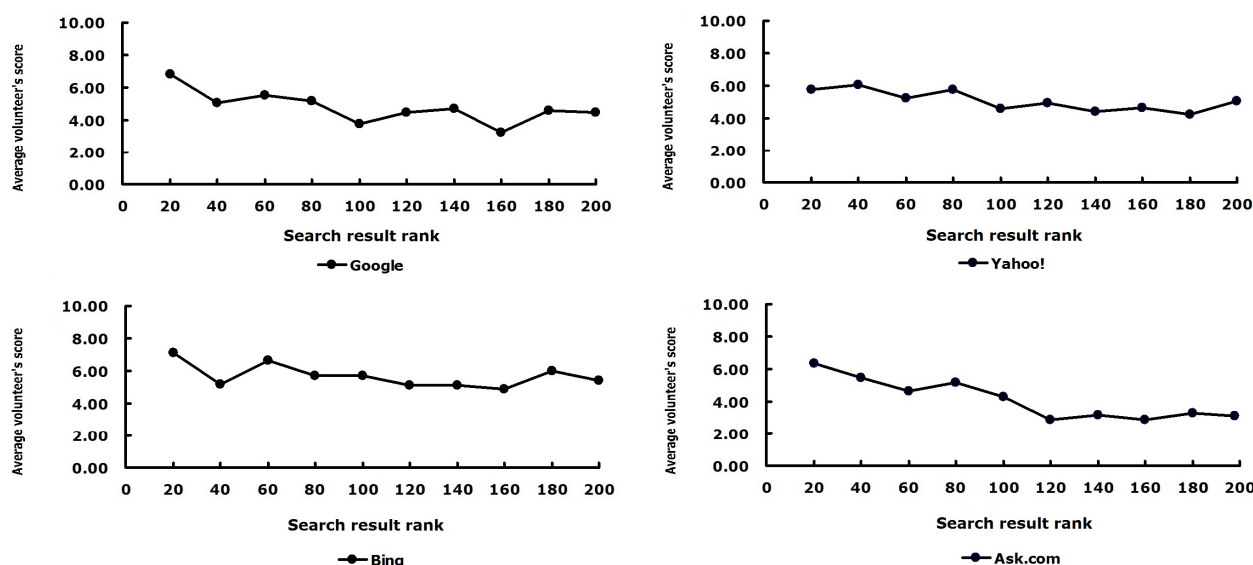
Figure 6. Distribution of websites with different scores in each search group (all eight volunteers' scores aggregated). (a) All search results, (b) top 20 search results.



(a)



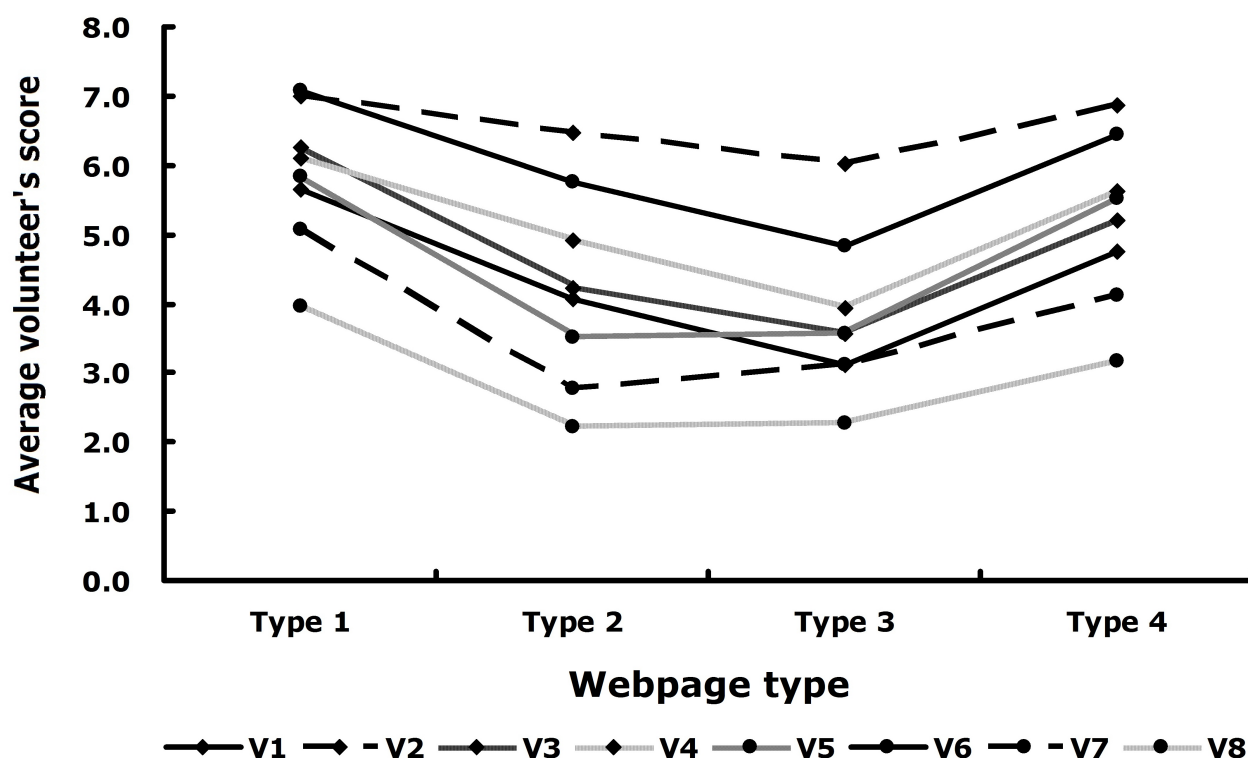
(b)

Figure 7. Volunteer's score versus search result rank (with a window size of 20) in each search group.

Difference Between Evaluations of Various Users

Volunteers had different user experiences and different schemes for grading webpages. Figure 8 plots each volunteer's average scores for the four website types. Volunteer 1, volunteer 2,

volunteer 3, volunteer 4, and volunteer 6 had similar tendencies, while volunteer 5, volunteer 7, and volunteer 8 had similar tendencies. Interestingly, the latter three volunteers were the youngest in the pool, although the sample size is too small to draw a conclusion on the effect of age.

Figure 8. Average volunteer's scores for the four webpage types. V1-8 = volunteers 1 to 8.

Keyword Frequencies in Search Results

Table 5 shows the occurrence in the webpages of *breast cancer* and the four other related keywords defined by SNOMED CT (*malignant tumor*, *neoplasm*, *sarcoma*, and *carcinoma*). The number of hits for the keyword *breast cancer* is much larger than that of any of the other four. The keyword *breast cancer*

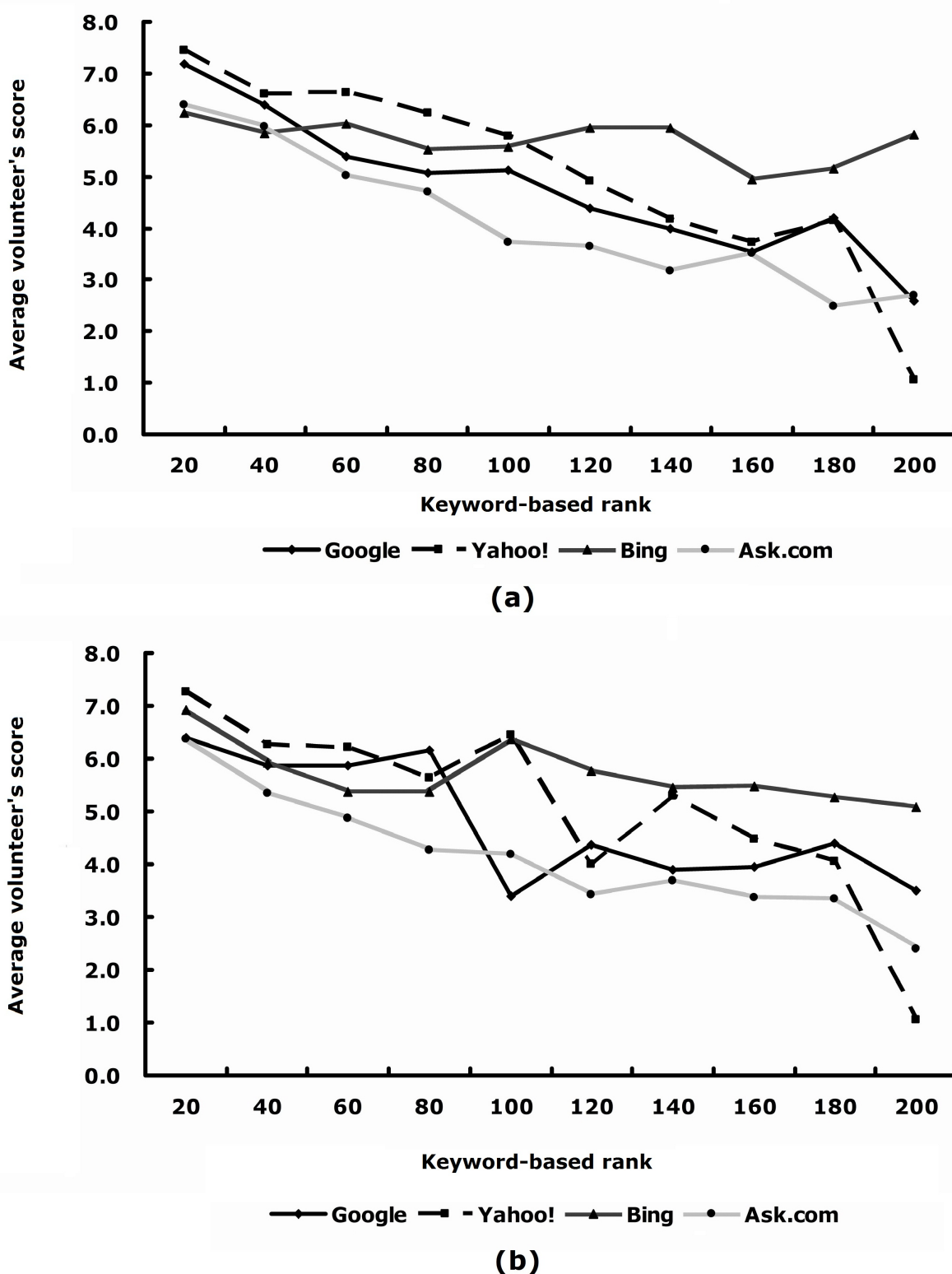
occurred among 91.71% (22,301/24,317) of the main page text corpus, and 88.53% (261,067/294,875) of the main and subpages text corpus. *Carcinoma* was the second most-frequent keyword, with an occurrence of 6.08% (1479/24,317) in the main page text corpus and of 6.87% (20,251/294,875) in the main and subpages text corpus.

Table 5. Occurrence of keywords in the main page text corpus (MP) and in the main and subpage text corpus (MSP).

Search engine	Breast cancer	Malignant tumor	Neoplasm	Sarcoma	Carcinoma
Google					
MP	4619	14	56	62	363
MSP	54,511	61	793	570	2629
Yahoo!					
MP	6423	28	77	45	399
MSP	70,471	128	1457	1085	5104
Bing					
MP	8171	23	80	80	548
MSP	89,577	143	2231	5355	10,397
Ask.com					
MP	3088	2	37	33	169
MSP	46,508	23	730	981	2121
Total					
MP	22,301	67	250	220	1479
MSP	261,067	355	5211	7991	20,251
Percentage					
MP	91.71%	0.3%	1.0%	0.90%	6.08%
MSP	88.53%	0.12%	1.77%	2.71%	6.87%

Figure 9 shows the average scores versus keyword-based reranking (as described in the subsection “Ranking based on selected keywords”) with an interval of 20 search results (top 1–20, ranks 21–40, ranks 41–60, etc) for each search group. As

the keyword-based ranking decreases, the average score also decreases, which indicates that the scores reflect an accurate representation of keywords.

Figure 9. Average volunteer's score versus keyword-based rank in (a) the main page text corpus and (b) the main and subpage text corpus.

Discussion

In this study, we compared and evaluated four major search engines for medical information search using various assessment criteria. The search results were significantly different between any two search engines. Yahoo! had the lowest performance scores, which does not correspond to its second ranking in the

market. A possible explanation of this discrepancy is that Yahoo! has been a major search engine for a long time and its other functions result in some good user experiences. Ask.com was not as good as Google and Bing. While Ask.com had a higher correlation between the scores and the ranking by the search engine than either Google or Bing, the volunteers scored its top hits lower, indicating a less useful status, as most users

browse just the top 10 to 20 hits. Google and Bing each have some merits. Google search results had higher validity and less redundancy. On the other hand, volunteers regarded the top hits of Bing as being more useful.

By studying the distribution of four types of webpages, we found that the search engines had different priorities for various categories of search results as shown in Figure 3 and Figure 4. Statistical results show that the content and type of search results from the Google and Yahoo! search engines were diversified with balanced content, and with a relatively low amount of advertisement content. In contrast, Bing had unbalanced search results, covering more breast cancer knowledge for the general population and professionals but fewer nonprofit organization websites and corporate websites for consumers. Bing's search results contained more advertisements than Google and Yahoo!. Ask.com also had unbalanced search results, emphasizing nonprofit organization websites for breast cancer patients and their families. Ask.com's search results had excessive advertisements, and its ability to control the advertisement information was the worst among the four engines, which might be related to the features of a metasearch engine [32].

Volunteers scored popular science and personal websites the highest, with an average score of 5.94; the second highest average score, 5.21, was for websites of public welfare organizations; and the third highest average score, 4.25, was for websites of universities, research institutes, hospitals, and governments. Volunteers gave the lowest score to corporation websites and advertisement websites, with an average of only 3.80. By interviewing and communicating with volunteers, we found that they had various Internet search habits. In most cases, users may go through the first few pages of the search results of a particular search engine. When the results are not good enough, they change the keywords to search again. They often choose additional keywords suggested by the search engines when they redo the search. Most users check only the search results listed on the first summary page. Some use different search engines for the same keyword search. Users who have little search experience often blindly trust the results from the search engines. They often believe that the top-ranked result is the best one. During the interviews, users also gave us some good suggestions for collecting data. For example, we provided them with standard websites, but did not provide them with a standard for low scores. Although the standards were authoritative, they could not accommodate various needs of different users; without a medical background, users may have a hard time evaluating professional websites.

Using search engines to obtain health and medical information is an effective method for most Internet users. Our study indicates that the four major search engines, Google, Yahoo!, Bing, and Ask.com, are all helpful to users in their health and medical searches. Thus, they are used and recommended by most consumers for obtaining medical information online. However, there is significant room for improvement, especially in getting more relevant and comprehensive information, as well as in ranking the websites according to their usefulness.

In this regard, there is no gold standard, and the various search engines each have their own merit, although Google and Bing are more advanced than others. Furthermore, the various search engines have different focuses on their search content. Hence, we suggest that users apply multiple search engines when looking for medical and health information online, instead of using only a single search engine.

Our study complements some earlier studies in evaluating Internet searching for health and medical information. Many of the previous studies emphasized the quality or reliability of the content [33]. In contrast, our study has a new focus—that is, usability—which more directly reflects the perspectives of actual users than most previous work. Our method is also significantly different from earlier approaches. For example, a previous usability study [34] was carried out in a usability laboratory. While conducting a study in a usability laboratory has some advantages, it also has some limitations. In particular, a usability laboratory is an artificial environment, and some participants may have different usage behavior especially under the time pressure of a study, as Eysenbach and Köhler [34] reported. Therefore, we used another method of usability testing: hallway testing. To best ensure that each participant sufficiently evaluated each webpage of the sample search results, we did not give a deadline for the evaluation and we only advised each participant to finish the evaluation as soon as possible to reduce the chances of possible changes and expirations of the websites (with the actual timeframes of the study, we believe such chances are negligible for any conclusion drawn in this paper). Some of our conclusions are in agreement with Eysenbach and Köhler [34]. For example, both studies found that consumers find information from popular science and personal websites and the websites of public welfare organizations acceptable. Nevertheless, our study sheds some new light on the usability of the various Internet search engines. Overall, our study adds significant value to the understanding of Internet searching for health and medical information.

In summary, our study may provide a useful analysis of medical information Internet searching and some helpful suggestions for improving the overall usability of health-related Internet searches. It provides some information to Internet users in terms of whether to use multiple search engines and how to use them. It also gives some informative data for Internet search engine developers to improve their search engine or to develop a medicine-specific search engine.

Our study also has some limitations. It had a limited sample size and used only one search keyword. Furthermore, the volunteers' backgrounds were relatively homogeneous, as they were all highly educated, usually in science. More large-scale studies with participants of diverse backgrounds are needed for conclusive results. While this study provides Internet users with an informative reference and some guidance for obtaining medical information online, we will conduct a much larger-scale study with more representative samples in the future. In particular, we will include cancer patients and their family members in the study.

Acknowledgments

The authors are grateful for the support of the National Natural Science Foundation of China (Grant No. 61073075), the Science-Technology Development Project of Jilin Province of China (Grant No. 201201048 and 20120730), and Basic Scientific Research from Jilin University (Grant No. 450060481090). We would also like to thank the anonymous reviewers of this paper for their helpful suggestions.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of the data and analyses. RN: rank of each result in the search list. The RN starts with 1001 in blue for Google results, 2001 in pink for Yahoo!, 3001 in green for Bing, and 4001 in red for Ask.com. Website name: the title of search result website. Website URL: the URL for search result website. GPR: Google PageRank Value (N means unavailable). Type: the four content types of search result websites. V1 to V8: the eight volunteers' scores of search results. AVG: the average score of a search result from all 8 volunteers. G, Y, B, A: the RN of the same website in the search results of Google, Yahoo!, Bing, and Ask.com, respectively. Main_Breast Cancer: frequency of keyword "Breast Cancer" in main page of search result. Sub_Breast Cancer: frequency of keyword "Breast Cancer" in main and subpage of search results combined.

[[XLS File \(Microsoft Excel File\), 616KB - jmir_v14i3e74_app1.xls](#)]

References

1. Taylor H, Leitman R. eHealth traffic critically dependent on search engines and portals. *Healthcare News* 2001;13:1-3 [[FREE Full text](#)]
2. Fox S. Health information online. Washington, DC: Pew Internet & American Life Project; 2005 May 17. URL: <http://www.pewinternet.org/Reports/2005/Health-Information-Online/1-Summary-of-Findings.aspx> [accessed 2011-09-22] [[WebCite Cache ID 61shFIIGc](#)]
3. Law MR, Mintzes B, Morgan SG. The sources and popularity of online drug information: an analysis of top search engine results and web page views. *Ann Pharmacother* 2011 Mar;45(3):350-356. [doi: [10.1345/aph.1P572](#)] [Medline: [21343404](#)]
4. Miller N, Lacroix EM, Backus JE. MEDLINEplus: building and maintaining the National Library of Medicine's consumer health Web service. *Bull Med Libr Assoc* 2000 Jan;88(1):11-17. [Medline: [10658959](#)]
5. Heilman JM, Kemmann E, Bonert M, Chatterjee A, Ragar B, Beards GM, et al. Wikipedia: a key tool for global public health promotion. *J Med Internet Res* 2011;13(1):e14 [[FREE Full text](#)] [doi: [10.2196/jmir.1589](#)] [Medline: [21282098](#)]
6. Greenwald R. . . . And a diagnostic test was performed. *N Engl J Med* 2005 Nov 10;353(19):2089-2090 [[FREE Full text](#)] [doi: [10.1056/NEJM200511103531923](#)] [Medline: [16282191](#)]
7. Jeannot JG, Scherer F, Pittet V, Burnand B, Vader JP. Use of the World Wide Web to implement clinical practice guidelines: a feasibility study. *J Med Internet Res* 2003 Jun;5(2):e12 [[FREE Full text](#)] [doi: [10.2196/jmir.5.2.e12](#)] [Medline: [12857668](#)]
8. Patel CO, Garg V, Khan SA. What do patients search for when seeking clinical trial information online? *AMIA Annu Symp Proc* 2010;2010:597-601. [Medline: [21347048](#)]
9. Spink A, Jansen BJ. *Web Search: Public Searching of the Web (Information Science and Knowledge Management)*. New York, NY: Springer; 2004.
10. Lee U, Liu Z, Cho J. Automatic identification of user goals in web search. In: *Proceedings. 2005 Presented at: 14th International World Wide Web Conference; May 10-14, 2005; Chiba, Japan.*
11. Broder A. A taxonomy of web search. *ACM SIGIR Forum* 2002;36(2):3-10. [doi: [10.1145/792550.792552](#)]
12. Steinbrook R. Searching for the right search--reaching the medical literature. *N Engl J Med* 2006 Jan 5;354(1):4-7. [doi: [10.1056/NEJMp058128](#)] [Medline: [16394296](#)]
13. 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](#)]
14. Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home. *BMJ* 1997 Jun 28;314(7098):1875-1879. [Medline: [9224132](#)]
15. Leighton HV, Srivastava J. First 20 precision among World Wide Web search services (search engines). *J Am Soc Inf Sci* 1999;50(10):870-881. [doi: [10.1002/\(SICI\)1097-4571\(1999\)50:10<870::AID-ASIA>3.0.CO;2-G](#)]
16. 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](#)]
17. Baeza-Yates R, Ribeiro-Neto B. *Modern Information Retrieval*. Reading, MA: Addison-Wesley Longman; 1999.

18. North American Consumer Technographics. Forrester Research. 2005 Jul. Consumer Technographics Q2 2005 North American Devices & Access Online Survey URL: <http://www.forrester.com/ER/Research/Survey/Excerpt/0,5449,430,00.html> [accessed 2011-09-22] [WebCite Cache ID 61sfARswm]
19. comScore, Inc. 2009 Oct 14. comScore Releases September 2009 U.S. Search Engine Rankings URL: [http://www.comscore.com/Press Events/Press Releases/2009/10/comScore Releases September 2009 U.S. Search Engine Rankings](http://www.comscore.com/Press%20Events/Press%20Releases/2009/10/comScore_Releases_September_2009_U.S._Search_Engine_Rankings) [accessed 2011-09-22] [WebCite Cache ID 61sfgTkAf]
20. Obama B. Whitehouse.gov. 2009 Sep 30. National Breast Cancer Awareness Month, 2009: A Proclamation URL: http://www.whitehouse.gov/the_press_office/Presidential-Proclamation-National-Breast-Cancer-Awareness-Month [accessed 2011-09-22] [WebCite Cache ID 61shN3FvY]
21. Wang L, Wang J, Wang M, Liang Y, Xu D. Use experience evaluation of Google search for obtaining medical knowledge: a case study. *Int J Data Min Bioinform* 2011;5(6):626-639. [Medline: [22295748](#)]
22. US National Library of Medicine. 2011 Mar 22. SNOMED Clinical Terms (SNOMED CT) URL: http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html [accessed 2011-09-22] [WebCite Cache ID 61siBcoo6]
23. Elkin PL, Brown SH, Husser CS, Bauer BA, Wahner-Roedler D, Rosenbloom ST, et al. Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists. *Mayo Clin Proc* 2006 Jun;81(6):741-748. [doi: [10.4065/81.6.741](#)] [Medline: [16770974](#)]
24. JExcelApi. Java Excel API: A Java to Read, Write, and Modify Excel Spreadsheets URL: <http://jexcelapi.sourceforge.net/> [accessed 2011-09-22] [WebCite Cache ID 61sg5wmPK]
25. Page L, Brin S, Motwani R, Winograd T. Stanford University InfoLab. Stanford, CA; 1999. The PageRank Citation Ranking: Bringing Order to the Web URL: <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf> [accessed 2012-04-25] [WebCite Cache ID 67Bo1rqvA]
26. Whatsmypagerank.com. 2009. What's My PageRank URL: <http://www.whatsmypagerank.com/> [accessed 2011-09-22] [WebCite Cache ID 61sgDxtRg]
27. Google PageRank Checker. PRChecker.info. 2011. Free Page Rank Checker for Your Web Site! URL: <http://www.prchecker.info/> [accessed 2011-09-22] [WebCite Cache ID 61sgK2mtG]
28. Nielsen J. Jakob Nielsen's Alertbox. 2000 Mar 19. Why You Only Need to Test With 5 Users URL: <http://www.useit.com/alertbox/20000319.html> [accessed 2011-09-22] [WebCite Cache ID 61sgOZB11]
29. Huang MH. Pausal behavior of end-users in online searching. *Inf Process Manag* 2003;39(3):425-444. [doi: [10.1016/S0306-4573\(02\)00040-7](#)]
30. Large A, Beheshti J, Rahman T. Gender differences in collaborative web searching behavior: an elementary school study. *Inf Process Manag* 2002;38(3):427-443. [doi: [10.1016/S0306-4573\(01\)00034-6](#)]
31. Bilal D, Kirby J. Differences and similarities in information seeking: children and adults as web users. *Inf Process Manag* 2002;38(5):649-670. [doi: [10.1016/S0306-4573\(01\)00057-7](#)]
32. Zhang J, Cheung C. Meta-search-engine feature analysis. *Online Inf Rev* 2003;27(6):433-441. [doi: [10.1108/14684520310510073](#)]
33. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 2002;287(20):2691-2700 [FREE Full text] [Medline: [12020305](#)]
34. 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. [Medline: [11884321](#)]

Abbreviations

SNOMED CT: Systematized Nomenclature of Medicine-Clinical Terms

Edited by G Eysenbach; submitted 23.09.11; peer-reviewed by XW Chen, J Powell, Z Cai, L Horn; comments to author 26.10.11; revised version received 26.12.11; accepted 17.01.12; published 16.05.12.

Please cite as:

Wang L, Wang J, Wang M, Li Y, Liang Y, Xu D
Using Internet Search Engines to Obtain Medical Information: A Comparative Study
J Med Internet Res 2012;14(3):e74
URL: <http://www.jmir.org/2012/3/e74/>
doi: [10.2196/jmir.1943](#)
PMID: [22672889](#)

Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Sensitivity and Predictive Value of 15 PubMed Search Strategies to Answer Clinical Questions Rated Against Full Systematic Reviews

Thomas Agoritsas^{1,2}, MD; Arnaud Merglen³, MD; Delphine S Courvoisier^{1,4}, PhD; Christophe Combescure¹, PhD; Nicolas Garin², MD; Arnaud Perrier², MD; Thomas V Perneger¹, MD, PhD

¹Division of Clinical Epidemiology, University Hospitals of Geneva, Geneva, Switzerland

²Division of General Internal Medicine, University Hospitals of Geneva, Geneva, Switzerland

³Division of General Pediatrics, Geneva Faculty of Medicine, University Hospitals of Geneva, Geneva, Switzerland

⁴Department of Psychology, Harvard University, Cambridge, MA, United States

Corresponding Author:

Thomas Agoritsas, MD

Division of Clinical Epidemiology

University Hospitals of Geneva

Gabrielle Perret-Gentil 6

Geneva, 1211 Geneva 14

Switzerland

Phone: 41 795534304

Fax: 41 223729035

Email: thomas.agoritsas@gmail.com

Abstract

Background: Clinicians perform searches in PubMed daily, but retrieving relevant studies is challenging due to the rapid expansion of medical knowledge. Little is known about the performance of search strategies when they are applied to answer specific clinical questions.

Objective: To compare the performance of 15 PubMed search strategies in retrieving relevant clinical trials on therapeutic interventions.

Methods: We used Cochrane systematic reviews to identify relevant trials for 30 clinical questions. Search terms were extracted from the abstract using a predefined procedure based on the population, interventions, comparison, outcomes (PICO) framework and combined into queries. We tested 15 search strategies that varied in their query (PIC or PICO), use of PubMed's Clinical Queries therapeutic filters (broad or narrow), search limits, and PubMed links to related articles. We assessed sensitivity (recall) and positive predictive value (precision) of each strategy on the first 2 PubMed pages (40 articles) and on the complete search output.

Results: The performance of the search strategies varied widely according to the clinical question. Unfiltered searches and those using the broad filter of Clinical Queries produced large outputs and retrieved few relevant articles within the first 2 pages, resulting in a median sensitivity of only 10%–25%. In contrast, all searches using the narrow filter performed significantly better, with a median sensitivity of about 50% (all $P < .001$ compared with unfiltered queries) and positive predictive values of 20%–30% ($P < .001$ compared with unfiltered queries). This benefit was consistent for most clinical questions. Searches based on related articles retrieved about a third of the relevant studies.

Conclusions: The Clinical Queries narrow filter, along with well-formulated queries based on the PICO framework, provided the greatest aid in retrieving relevant clinical trials within the 2 first PubMed pages. These results can help clinicians apply effective strategies to answer their questions at the point of care.

(*J Med Internet Res* 2012;14(3):e85) doi:[10.2196/jmir.2021](https://doi.org/10.2196/jmir.2021)

KEYWORDS

Evidence-based medicine; information retrieval; medical literature; search strategy; PubMed; Medline; clinical queries; search filters; sensitivity; recall; positive predictive value; precision

Introduction

Searching the literature for evidence has become a central skill in clinical practice [1]. Physicians' information needs are considerable [2-4], and evidence-based decisions often require the identification and appraisal of current research findings [5]. As Glasziou et al commented, "the use of search engines is now as essential as the stethoscope" [6]. However, retrieving relevant information has also become increasingly challenging, given the rapid expansion of medical knowledge: PubMed now comprises more than 20 million citations [7], and 2000–4000 more are added every day [6], including 75 clinical trials and 11 systematic reviews [8].

In the past decade, several solutions have been implemented to improve access to current research [9]. Preappraised resources, such as evidence-based medicine journals, or Web-based summaries such as UpToDate [10], have been adopted by many clinicians [11,12]. But these resources are limited by delayed processing [9,13], insufficient coverage [9,14], or cost [15,16]. Thus, PubMed remains the most popular search engine used to retrieve original studies [17-20], either alone or as a complement to reappraised resources [15,18].

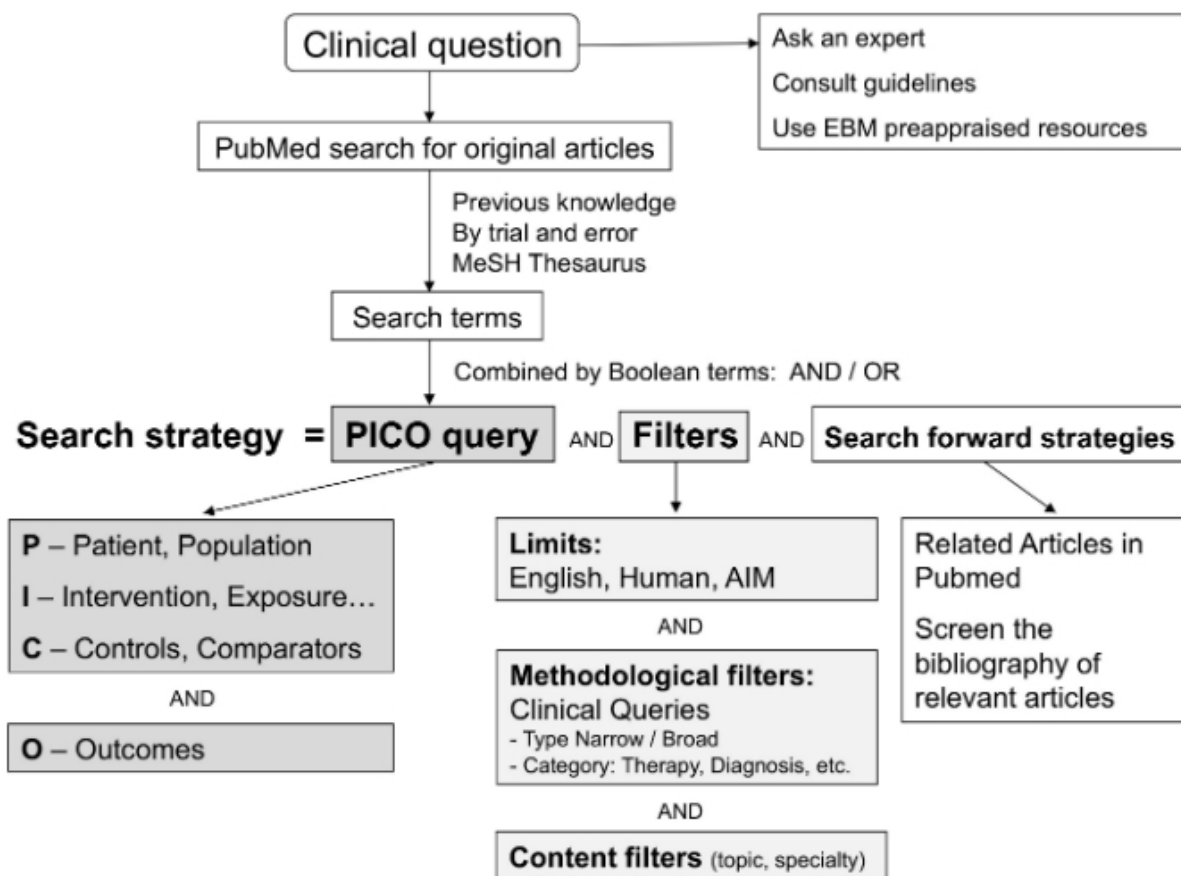
However, searching PubMed is not as intuitive as searching other commonly used engines such as Google. While clinicians often perform short unstructured queries of 2–3 terms [21], these tend to produce large and diluted outputs. More efficient search strategies can be proposed (Figure 1), using existing search tools and based on expert recommendations. First, the clinical question is translated into search terms that are combined into a query using Boolean operators (eg, OR, AND) [22,23]. The use of the population, interventions, comparison, outcomes (PICO) framework helps formulate more precise queries that

combine search terms for these four factors [22,24]. The size of the output can be further reduced using limits [25] or methodological filters, such as PubMed's Clinical Queries [26], designed to retrieve high-quality randomized controlled trials for therapeutic interventions [27]. Finally, additional strategies, such as the use of the related articles link in PubMed, can identify studies based on a relevant article that was initially found (Figure 1 [28]).

Evidence regarding the performance of search strategies to answer clinical questions is scarce. A few small studies have assessed the impact of giving clinicians search tutorials on the retrieval of specific sets of articles identified by experts, but search strategies could not be compared, as clinicians were free to use their own informal strategies [29,30]. Other studies have reported on the performances of search filters in retrieving high-quality clinical trials, but these filters were assessed independently of user queries [31] or clinical questions [27]. To our knowledge, there is little evidence to help clinicians organize their searches and combine existing search tools into effective strategies that are applicable at the point of care.

In this study, we compared the performance of 15 PubMed search strategies in retrieving relevant clinical trials, as identified by high-quality systematic reviews on specific clinical questions. We devised these 15 search strategies by choosing and combining the following search components that are easily applicable at the point of care: formulation of queries using the PICO framework, use of Clinical Queries therapeutic filters (broad or narrow), use of several search limits, and use of PubMed links to related articles. Our aim was to identify search components and tools that would most likely help clinicians answer questions on therapeutic interventions at the point of care.

Figure 1. Framework for PubMed search strategies to answer clinical questions. AIM = Abridged Index Medicus (set of 119 core clinical journal titles [28]), EBM = evidence-based medicine, MeSH = Medical Subject Headings (a controlled vocabulary used for indexing articles in PubMed), PICO = population, interventions, comparison, outcomes.



Methods

Sample of Systematic Reviews for the Identification of Relevant Articles

For a specific clinical question, a well-conducted systematic review is considered the gold standard for the identification of all relevant articles [31-33], since it implies a systematic and comprehensive search, as well as an appraisal of the articles' scientific rigor. Moreover, recent systematic reviews address questions of clinical interest for which relative uncertainty exists and are thus a good proxy of the questions that may arise in real clinical practice.

We selected 30 clinical questions from a wide spectrum of topics of general interest, on the basis of 30 recent Cochrane systematic reviews (Table 1, Multimedia Appendix 1). We favored Cochrane reviews because of their well-established quality and methodological rigor [34-36], and focused on reviews that included randomized controlled trials, as these constitute the highest level of evidence for questions about therapeutic interventions. In February 2010, we searched the Cochrane Database of Systematic Reviews [37] for all reviews on therapeutic interventions published since the beginning of 2010. To focus on the most recent reviews, or reviews that were recently updated, we restricted the search to citations with the

following record status: new review, new search, or conclusion changed. Of the 287 potentially eligible reviews, we excluded those with fewer than 4 studies ($n = 113$), those that were less relevant for a general audience ($n = 82$), and those with composite questions that would result in complex queries for clinical practice ($n = 62$). All clinical trials included in the remaining 30 reviews that were also retrievable in PubMed defined the subset of relevant studies used in our assessment.

Extraction of Search Terms and Formulation of PICO Query

We selected search terms using a predefined procedure that we determined before applying the searches strategies. Following the PICO framework (Figure 1), we categorized the exact wording of the objective and selection criteria of each review's abstract into four sections for the population, interventions, comparison, and outcomes [22,24]. When we could find no information in the abstract for one of these categories, we retrieved the relevant keywords from the methods section (eg, type of intervention or type of outcome). Then for each PICO category, we extracted the smallest set of search terms that best expressed the clinical question. This predefined procedure was performed by 2 assessors (TA and AM), both trained in clinical epidemiology and evidence-based medicine. In a learning phase, together they extracted search terms from a pilot sample of 10 systematic reviews. Then they independently extracted PICO

search terms from the 30 reviews of the study sample and agreed on the smallest set of terms that best expressed the clinical questions. They resolved discrepancies by consensus. Finally, all coauthors, including experienced clinicians, approved the retained search terms. For half of the reviews, PICO's comparison component shared common terms with interventions (eg, composite vs single intervention), and in these cases no term was retained for comparison. Examples of search term extraction are shown in [Multimedia Appendix 2](#).

To obtain the final search query ([Figure 1](#)), search terms were combined with the Boolean operator OR within each PICO category, so as to increase sensitivity. Then all categories were combined with AND to retrieve only the citations matching all PICO elements [22-24]. For example, for a review on the effects of oral mucolytic agents in adults with stable chronic bronchitis or chronic obstructive pulmonary disease (COPD) [38], the final full PICO query was *((chronic bronchitis) OR COPD) AND mucolytics AND placebo AND exacerbations*. Search terms consisting of groups of words (eg, chronic bronchitis) were put into parentheses, not in quotes. The detailed wording of all 30 queries can be found in [Multimedia Appendix 3](#).

Design of Search Strategies

The purpose and context of the search determine the choice of search strategy. Our objective was to identify search components and tools that could help clinicians build more effective strategies to answer questions at the point of care. Therefore, selecting sophisticated strategies used for performing systematic reviews [39], such as the Cochrane Highly Sensitive Search Strategy [40], would have been inappropriate. Since there is very little evidence on strategies adapted to our purpose, we devised our 15 search strategies ([Table 2](#), [Figure 1](#)) based on the few expert recommendations available [22-24,26,27,29-31]. We favored search components and search tools that are applicable and easy to use in clinical practice.

We thus obtained strategies by varying the following parameters. First, the search query was either the full PICO query or a truncated PIC query that did not include terms of outcomes. Although the full PICO is usually recommended for clinicians to help them formulate more precise searches [22-24], such strategies may also miss relevant articles [40], as the outcomes are less often mentioned in the abstract or assigned Medical Subject Headings (MeSH) in PubMed. Moreover, formulating shorter PIC queries may be less burdensome in practice and may be more appropriate for clinicians who are interested in many potential outcomes. We entered search terms using PubMed's automatic term mapping, which systematically searches each term in the MeSH thesaurus, as well as the "[all fields]" tag ([Multimedia Appendix 2](#)). We did not test strategies using the restriction to MeSH terms, as few clinicians manage to perform the necessary steps for this restriction in practice [21], and this proportion hardly increases after medical residents take a search tutorial [29].

Second, we combined these queries with either the broad or narrow filter of the Clinical Queries for therapeutic interventions. We chose Clinical Queries over other filters meant to improve the retrieval of high-quality randomized controlled trials [39] because Clinical Queries were specifically designed

to help clinicians answer their questions on therapeutic interventions and are implemented in PubMed [27]. To use these filters, clinicians simply need to enter their search terms on PubMed's Clinical Queries page [26]. After taking a search tutorial, about 80% of clinicians adopted the use of Clinical Queries [29]. Third, we repeated these strategies adding search limits [25] that restricted the searches to citations in English and research on humans. PICO searches with human and English limits were repeated with a further restriction to a set of 119 core medical journals listed in the Abridged Index Medicus (AIM) [28]. We did not further add the limit "clinical trials [pt]," as it would have been redundant with the Clinical Queries filters.

Finally, the last 3 searches tested the PubMed link to related articles, a search tool that identifies content similarity in the title, abstract, and index terms [41]. To use this tool in practice, clinicians must identify a citation they consider to be potentially relevant in a first search output (ie, a citation that best corresponds to their PICO question) and click on its link to related articles. They can then scan a new search output with citations ranked by content similarity with the initial citation. To test this strategy, 2 assessors (TA and AM) screened the first page of search #11 (PICO, therapy narrow filter, limited to human studies in English), a search strategy that yielded short outputs ([Table 2](#)), and identified by consensus the 3 citations that were closest to the corresponding PICO question. They performed this selection being blinded to the citations included in the review, so as to correspond to the situation that clinicians encounter in practice. Search performance was assessed on each output of these 3 citations' related articles. We performed this strategy based on 3 citations, instead of 1, to minimize the impact of the subjective component of the choice of initial citation.

Analysis of Search Performance

We applied all 15 strategies for each of the 30 clinical questions (450 searches in total). We restricted each search to the date when the corresponding review was assessed as up-to-date, so that the time frame of the search was the same as its corresponding review. For each search we collected the number of articles in the output, the number of gold standard articles retrieved, and their position in the output. Then we computed the sensitivity (also called recall: the proportion of relevant papers that were retrieved) of each search and its positive predictive value ([PPV], also called precision: the proportion of the output corresponding to relevant papers), as defined in [Figure 2](#) [27,31,33].

Because most physicians screen only 2 pages of the PubMed output, or at most 40 items [19,21], our primary outcomes for search performance were sensitivity and PPV for this cut-off. However, we also analyzed overall sensitivity and PPV on the full search output and further explored how these properties varied according to the cumulative number of items screened in the output.

We summarized the performance of the search strategies over the 30 clinical questions using nonparametric statistics and box plots. Filtered versus unfiltered searches were compared using the Wilcoxon signed rank tests (paired). Finally, we explored

whether retrieval performances differed according to the characteristics of the corresponding review. All analyses were performed using R 2.12.1 software [42].

Figure 2. Definitions of the sensitivity and positive predictive value (PPV) of search strategies, when using a systematic review as the gold standard.

^a We included only the studies published before the date when the corresponding systematic review was assessed as up-to-date. ^b We excluded from this total the studies that are not indexed in PubMed, as they cannot be retrieved by any search strategy. ^c This total is unknown, because it is not limited to the studies that were explicitly excluded from the systematic review. ^d When only a limited number of items in the output are screened (eg, 40 items, or 2 pages of PubMed's output), then practically this number becomes the true denominator of PPV.

Search strategy	Studies retrieved ^a	Relevant studies (included in the systematic review)	Nonrelevant studies (not included in the systematic review)	$x + z$ # items in the output or # items that are screened ^d
		x	z	
	Studies not retrieved	y	unknown	unknown
		$x + y$ ^b	unknown ^c	

$$\text{Sensitivity (or recall)} = \frac{\text{Relevant studies retrieved by the search}}{\text{Relevant studies included in the systematic review}} = \frac{x}{x + y}$$

$$\text{PPV (or precision)} = \frac{\text{Relevant studies retrieved by the search}}{\text{Number of items generated (or screened ^d) in the output}} = \frac{x}{x + z}$$

$$\text{NNR (Number Needed to Read to find a relevant article in the output)} = \frac{1}{\text{PPV}}$$

Results

Characteristics of Systematic Reviews and Corresponding PICO Queries

The 30 systematic reviews addressed a broad pattern of clinical topics and were produced by 15 different Cochrane groups, with a range of 1–4 reviews per group. Of these, 13 reviews (43%) had their record status labeled as new review, and 13 others (43%) as new search with no change to the conclusion and 4 (13%) with conclusion changed. They included a median of 15 studies (interquartile range [IQR] 7–23, range 5–49), of which 85% (IQR 66%–100%) were retrievable in PubMed (Table 1). This led to a median of 12 relevant clinical trials per review

(IQR 7–18) that we considered to be the gold standard for each clinical question.

The predefined extraction procedure of search terms from the reviews' abstracts led to a median of 2 terms (IQR 1–2) for population 1 term (IQR 1–2) for interventions, 1 term (IQR 0–1) for comparison, and 2.5 search terms (IQR 2–3) for outcomes. The required number of search terms was variable, as these terms were strictly tailored to the reviews' questions. Indeed, these were sometimes more complex or consisted of grouping of words (eg, generalized onset tonic-clonic seizures). Overall, queries resulted in a median of 4 terms (IQR 3–6) for the PIC query and 7 terms (IQR 5–8) for the PICO query (for detailed wording, see Multimedia Appendix 3).

Table 1. Comparative performance assessed on the full search output (A) and on 2-page output (B) of 2 selected search strategies—one unfiltered (S1) and one using Clinical Queries narrow filter (S4)—to answer 30 clinical questions rated against 30 Cochrane systematic reviews.

Rev. No.	Review title	No. studies incl. in re-view	n (%) retrievable in PubMed	A. Search performance for the full output				B. Search performance for a 2-page output (maximum 40 items)			
				Sensitivity (%)		PPV ^a (%)		Sensitivity (%)		PPV (%)	
				S1 ^b	S4 ^c	S1	S4	S1	S4	S1	S4
1	Carbamazepine versus phenytoin monotherapy for epilepsy	10	8 (80)	37.5	37.5	0.1	11.1	0.0	37.5	0.0	11.1
2	Chest physiotherapy for pneumonia in adults	6	5 (83)	100.0	80.0	1.9	12.5	0.0	80.0	0.0	12.5
3	Epinephrine injection versus epinephrine injection and a second endoscopic method in high risk bleeding ulcers	18	15 (83)	100.0	100.0	7.9	19.5	6.7	33.3	2.5	12.5
4	Extracranial-intracranial arterial bypass surgery for occlusive carotid artery disease	21	18 (86)	38.9	5.6	4.1	14.3	0.0	5.6	0.0	14.3
5	Fluticasone versus 'extrafine' HFA-beclomethasone dipropionate for chronic asthma in adults and children	9	8 (89)	100.0	100.0	18.6	42.1	100.0	100.0	20.0	42.1
6	Influenza vaccination for healthcare workers who work with the elderly	5	5 (100)	100.0	80.0	1.5	16.7	20.0	80.0	2.5	16.7
7	Mucolytic agents for chronic bronchitis or chronic obstructive pulmonary disease	29	18 (62)	83.3	66.7	11.5	18.8	22.2	38.9	10.0	17.5
8	Neuraminidase inhibitors for preventing and treating influenza in healthy adults	20	20 (100)	90.0	90.0	22.8	35.3	15.0	70.0	7.5	35.0
9	Adenoidectomy for otitis media in children	14	14 (100)	100.0	100.0	2.3	22.2	7.1	64.3	2.5	22.5
10	Antibiotics and antiseptics for venous leg ulcers	25	13 (52)	84.6	76.9	6.4	43.5	15.4	76.9	5.0	43.5
11	Antithyroid drug regimen for treating Graves' hyperthyroidism	27	23 (85)	82.6	78.3	4.9	36.7	0.0	69.6	0.0	40.0
12	Artesunate versus quinine for treating severe malaria	6	4 (67)	100.0	100.0	8.3	33.3	50.0	100.0	5.0	33.3

Rev. No.	Review title	No. studies incl. in review	n (%) retrievable in PubMed	A. Search performance for the full output				B. Search performance for a 2-page output (maximum 40 items)			
				Sensitivity (%)		PPV ^a (%)		Sensitivity (%)		PPV (%)	
				S1 ^b	S4 ^c	S1	S4	S1	S4	S1	S4
13	Bed rest for acute uncomplicated myocardial infarction	17	11 (65)	54.5	27.3	19.4	60.0	54.5	27.3	19.4	60.0
14	Benzodiazepines for the relief of breathlessness in advanced malignant and non-malignant diseases in adults	7	4 (57)	75.0	50.0	10.7	40.0	75.0	50.0	10.7	40.0
15	Blood pressure lowering efficacy of beta-blockers as second-line therapy for primary hypertension	20	11 (55)	72.7	63.6	1.2	2.3	0.0	0.0	0.0	0.0
16	Blood pressure lowering efficacy of potassium-sparing diuretics for primary hypertension	8	7 (88)	100.0	71.4	1.2	2.5	0.0	0.0	0.0	0.0
17	Caffeine for asthma	7	7 (100)	85.7	71.4	66.7	83.3	85.7	71.4	66.7	83.3
18	Continuous subcutaneous insulin infusion (CSII) versus multiple insulin injections for type 1 diabetes mellitus	34	27 (79)	66.7	55.6	9.9	44.1	7.4	55.6	5.0	44.1
19	Effect of cyclosporine on blood pressure	17	17 (100)	47.1	47.1	9.3	11.9	17.6	23.5	7.5	10.0
20	Enteral versus parenteral nutrition for acute pancreatitis	8	8 (100)	100.0	100.0	3.9	30.8	12.5	100.0	2.5	30.8
21	Exercises for prevention of recurrences of low-back pain	13	13 (100)	84.6	84.6	0.5	2.7	0.0	0.0	0.0	0.0
22	Home-based versus centre-based cardiac rehabilitation	22	17 (77)	58.8	47.1	7.8	15.4	29.4	47.1	12.5	20.0
23	Immediate-release versus controlled-release carbamazepine in the treatment of epilepsy	10	10 (100)	80.0	80.0	11.8	34.8	30.0	80.0	7.5	34.8
24	Proton pump inhibitor treatment initiated prior to endoscopic diagnosis in upper gastrointestinal bleeding	6	3 (50)	33.3	33.3	0.9	14.3	0.0	33.3	0.0	14.3

Rev. No.	Review title	No. studies incl. in re-view	n (%) retrievable in PubMed	A. Search performance for the full output				B. Search performance for a 2-page output (maximum 40 items)			
				Sensitivity (%)		PPV ^a (%)		Sensitivity (%)		PPV (%)	
				S1 ^b	S4 ^c	S1	S4	S1	S4	S1	S4
25	Rectal 5-aminosalicylic acid for induction of remission in ulcerative colitis	38	33 (87)	90.9	84.8	10.3	32.6	12.1	39.4	10.0	32.5
26	Regular treatment with salmeterol for chronic asthma: serious adverse events	49	30 (61)	43.3	40.0	7.6	18.8	0.0	16.7	0.0	12.5
27	Rifabutin for treating pulmonary tuberculosis	5	4(80)	100.0	100.0	2.7	33.3	0.0	100.0	0.0	33.3
28	Serotonin receptor antagonists for highly emetogenic chemotherapy in adults	16	15(94)	46.7	46.7	2.2	7.3	6.7	13.3	2.5	5.0
29	Short-term treatment with proton pump inhibitors, H2-receptor antagonists and prokinetics for gastro-oesophageal reflux disease-like symptoms and endoscopy negative reflux disease	45	38 (80)	57.9	55.3	0.7	4.4	0.0	0.0	0.0	0.0
30	Therapeutic ultrasound for osteoarthritis of the knee or hip	5	5 (100)	100.0	100.0	9.4	29.4	100.0	100.0	12.5	29.4

^a Positive predictive value (ie, precision).

^b Unfiltered search No. 1 corresponding to a single population, interventions, comparison (PIC) query, without any filters or limits (see also Table 2).

^c Filtered search No. 4 corresponding to a PIC query, combined with the narrow therapy filter of PubMed's Clinical Queries (see also Table 2).

Performance of Search Strategies

We observed important differences in the sizes of the output across search strategies (Table 2). Unfiltered PIC queries resulted in the largest outputs, with a median of 173 items, while the PICO queries halved the output size. The use of the Clinical Queries broad filter reduced the output by about 20%, and use of the narrow filter reduced it by about 80%. In contrast, searches based on related articles typically retrieved hundreds of articles.

The sensitivity and the positive predictive value (PPV) were also highly variable within each search strategy (Figure 3). When the full outputs were screened for relevant studies, about 85% were detected by PIC queries and 69% by PICO queries (Figure 3A). Overall sensitivity remained comparable when we used Clinical Queries filters (with or without limits), although the Clinical Queries narrow filter was associated with slightly lower overall sensitivities (Table 1A). In contrast, the use of the AIM limit systematically lowered sensitivity to about 15% (Figure 3A). The overall sensitivity of searches based on related

articles was also extremely variable, with a median of about 60%.

When the screening of relevant articles was limited to the first 2 pages (ie, 40 articles), sensitivity dropped to 10% for unfiltered searches and PPV decreased to 2.5%–10% (Figure 3B). In contrast, all search strategies that used the Clinical Queries narrow filter (S4, S5, S10, and S11 in Table 2) showed significantly higher sensitivities, overall around 50% ($P < .001$ when compared with their corresponding unfiltered queries) and higher PPV values, between 20% and 30% ($P < .001$) within the first 2 pages of the output. When looking at the 30 questions individually, adding a Clinical Queries narrow filter to a query increased the sensitivity of the search for 21 questions (70%), kept it stable for 6 (20%), and decreased it for 3 (10%), whereas PPV increased in 27 questions (90%) (Table 1B). Moreover, adding a Clinical Queries narrow filter to a PIC query reduced the risk of finding no relevant articles in the first 2 pages from 11 questions (37%) to 4 questions (13%). The broad Clinical Queries filter did not improve search performance; nor did searches using related articles and the additional use of limits.

Overall, PICO queries had slightly higher performances than PIC queries, although the differences were not statistically significant.

The exploration of search performance according to the number of articles screened in the output showed that sensitivity rose gradually up to about 100 items in unfiltered searches (Figure 4). In contrast, the sensitivity of searches using the Clinical

Queries narrow filter rose much more steeply, peaking earlier at 50 to 60 articles screened. Finally, we found no significant association between search performance and characteristics of the reviews, namely the absolute number of relevant articles included, the number of PICO search terms required to summarize its question, or the presence of mortality as an outcome (data not shown).

Table 2. Searches strategies: description, number of hits, and performances over the first 2 pages of PubMed output.

Strategy No.	Search strategy			No. of hits in output		Performances for an output of 2 pages (maximum 40 items)				
	Query	Clinical Queries	Limits	Median	IQR ^a	Sensitivity (%)		PPV ^b (%)		NNR ^c
						Median	IQR	Median	IQR	Median
S1	PIC ^d	NA ^e	NA	173	79–322	9.8	0.0–29.4	2.5	0.0–10.0	40
S2	PIC	Therapy, broad	NA	126	66–276	14.6	0.0–30.0	5.0	0.0–10.0	20
S3	PIC	Therapy, broad	English, human	97	59–229	17.6	0.0–33.3	5.0	0.0–12.5	20
S4	PIC	Therapy, narrow	NA	33	17–67	48.5	23.5–80.0	21.3	12.5–35.0	5
S5	PIC	Therapy, narrow	English, human	31	14–67	52.8	23.5–80.0	23.8	12.5–36.4	4
S6	PICO ^d	NA	NA	91	36–179	17.9	4.3–60.0	6.3	2.5–15.0	16
S7	PICO	Therapy, broad	NA	75	33–165	26.1	9.1–75.0	8.8	5.0–20.0	11
S8	PICO	Therapy, broad	English, human	62	28–138	29.6	9.1–75.0	11.3	5.0–20.8	9
S9	PICO	Therapy, broad	English, human, AIM ^f	11	5–24	15.5	11.1–40.0	20.0	14.3–33.3	5
S10	PICO	Therapy, narrow	NA	22	13–51	54.7	27.3–78.6	32.1	14.3–50.0	3
S11	PICO	Therapy, narrow	English, human	20	12–50	54.7	27.3–78.6	32.8	15.0–50.0	3
S12	PICO	Therapy, narrow	English, human, AIM	5	3–13	15.5	10.5–33.3	50.0	23.1–56.2	2
S13	Related #1 ^g	NA	NA	350	204–599	39.7	20.0–50.0	10.0	5.0–15.0	10
S14	Related #2 ^g	NA	NA	340	138–484	37.9	18.4–62.5	10.0	5.0–17.5	10
S15	Related #3 ^g	NA	NA	305	167–558	37.5	18.5–50.0	7.5	5.0–17.5	13

^a Interquartile range.

^b Positive predictive value (ie, precision).

^c Number of items needed to read to find a relevant article in the screened output; equal to 1/PPV of the search.

^d Population, interventions, comparison, (outcomes).

^e Not applied.

^f Abridged Index Medicus (set of 119 core clinical journals [28]).

^g From the output of search S11, related articles were searched for the 3 articles whose title was closest to the PICO query.

Figure 3. Compared sensitivity and positive predictive value (PPV) of the 15 search strategies (S) tested, for (A) the full search output and (B) the first 2 PubMed pages (40 articles). CQ = Clinical Queries, PIC(O) = population, interventions, comparison, (outcomes).

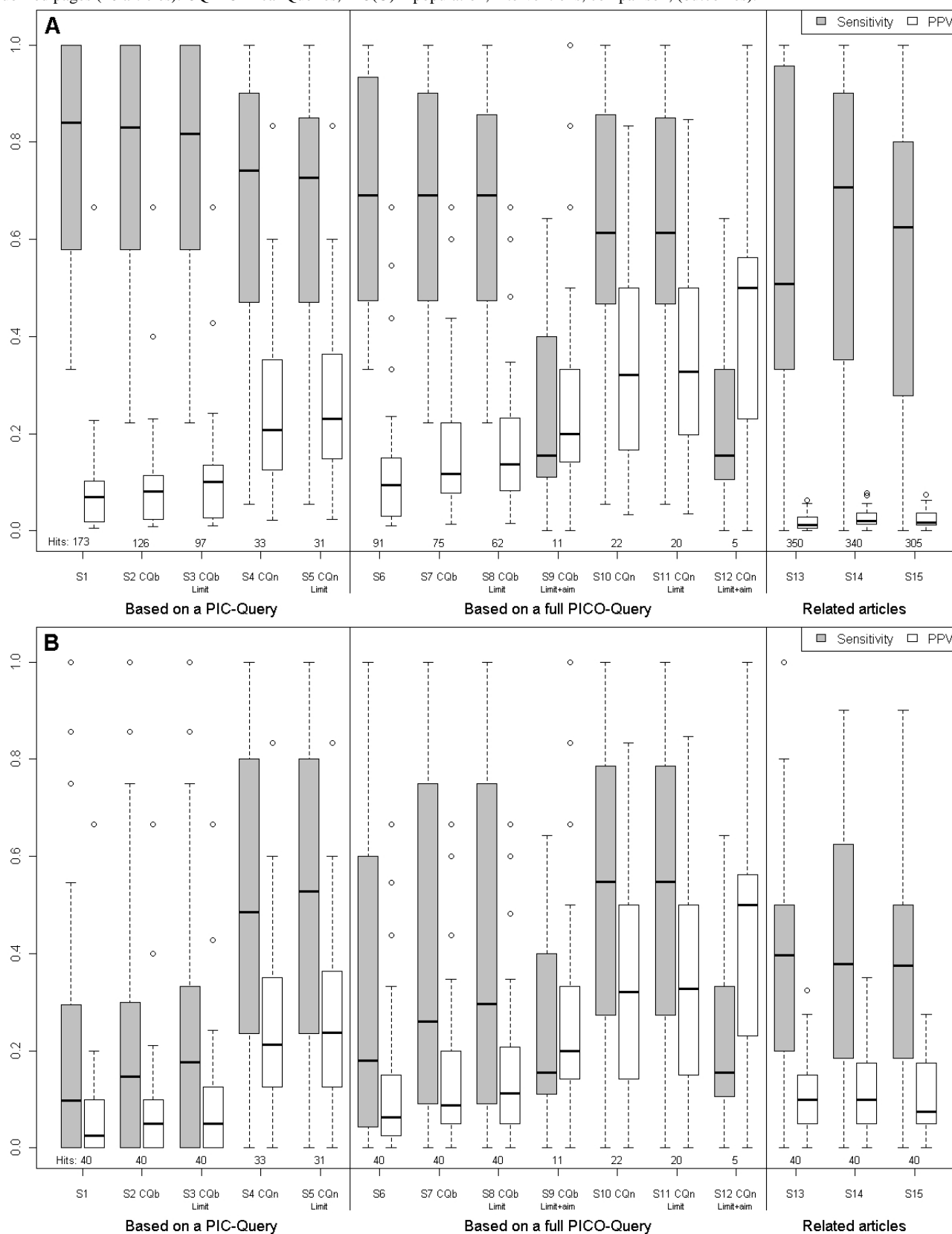
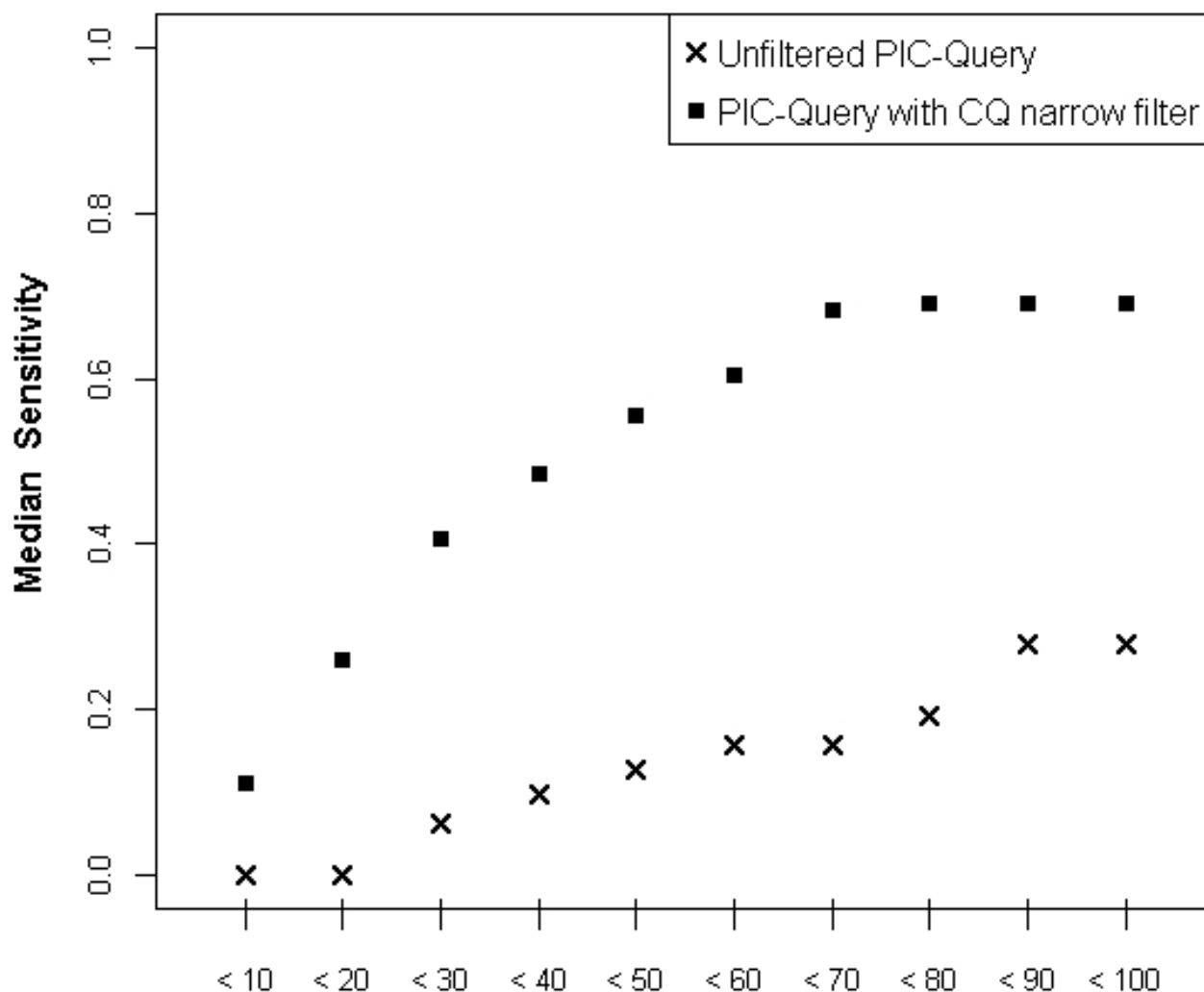


Figure 4. Evolution of median sensitivity of the search according to the number of studies screened in the output for the unfiltered population, interventions, comparison (PIC) query (S1) and the PIC query filtered by the Clinical Queries (CQ) narrow filter (S4).



Discussion

The performance of the search strategies tested was highly variable according to the clinical question. None of the 15 strategies showed a consistently high sensitivity in retrieving relevant articles identified by systematic reviews. Even the best strategies had a sensitivity of about 50%, with a range from 0% to 100% across the 30 questions. Therefore, clinicians who perform a specific PubMed search cannot foresee to what extent their search will be successful.

Nevertheless, on average, some search strategies were more successful than others. Unfiltered searches based on PIC or PICO queries, as well as those using the broad Clinical Queries filter, produced large outputs and retrieved few relevant articles within the first 2 pages. In contrast, searches that used the Clinical Queries narrow filter had a significantly higher sensitivity (around 50%–55%) and a higher PPV (around 20%–30%). These improved performances were observed for most clinical questions. Additional use of limits had only a marginal effect, except for the limit to core medical journals, which actually reduced sensitivity. Finally, searches based on related articles retrieved about a third of relevant studies in the first 2 pages. However, their PPV remained low, despite the

ranking of citations by content similarity, because they also retrieved hundreds of nonrelevant articles. This may be because related articles will also retrieve studies that are not randomized trials, in contrast with Clinical Queries therapeutic filters.

Clinicians who use PubMed at the point of care favor short queries of 2–3 terms, without any limits or filters, and generally screen only the first 2 pages of the output (ie, 20–40 items) [19,31]. Our results clearly show that such strategies will miss many relevant articles. Although these strategies may reach high overall sensitivities, relevant citations will be scattered over many pages that will not be screened in real life. To increase the density of relevant articles in the output (ie, the PPV), we recommend the use of the PIC(O) framework. Such queries including 5–8 terms kept a satisfactory overall sensitivity of about 85% (Figure 3A). However, when only the first 2 pages of the output were scanned, which is what usually occurs in clinical practice [21], this sensitivity dropped to about 10%, with an even lower PPV of 2.5% (corresponding to a median of 1 relevant article per screened output).

To improve search performance within readable outputs, an important finding is the usefulness of the Clinical Queries narrow therapy filter in identifying relevant studies in the first 2 pages of PubMed output. Clinical Queries filters were

designed by Haynes et al [27] based on a hand search of 161 clinical journals that identified methodologically sound randomized controlled trials. When applied to the whole Medline database, the Clinical Queries broad filter retained 99% of hand-selected studies, while the narrow filter retained 93% [27]. Hoogendam et al found similar results [31]. Our study confirmed the high sensitivity of the Clinical Queries filters when applied to specific clinical questions. A surprising finding was the increased sensitivity of the narrow filter compared with the broad filter when we screened only a realistic portion of the search output; this was because the narrow filter concentrated relevant articles on the first pages of the output. In other words, sensitivity, assessed on 2 pages of output, was better as PPV improved.

Based on our results we recommend that clinicians at the point of care start their PubMed searches by formulating a PIC query combined with a narrow filter, and then adapt it according to the output size and relevancy of the first items screened. If this strategy retrieves too many irrelevant citations, search terms describing the outcome (full PICO query) and further search limits can be added. In the opposite case, filters can be deactivated and sensitivity can be gained by adding search terms or by using the related articles link from the first relevant studies found. Finally, a general recommendation would be to screen beyond the first page of the output, as even the most efficient strategies require at least 50 to 60 items to approach their maximum sensitivity (Figure 4).

Study Limitations and Strengths

The main limitation of our study is that we assessed search strategies based solely on their retrieval performance. The searches were not performed by clinicians at the point of care, so we did not capture the iterative process of searching PubMed, based on trial and error [43,44], and could not assess whether relevant articles would be identified in practice. Furthermore, we did not examine the impact of search strategies on medical decisions and patients' outcomes. However, performing searches that retrieve relevant articles is a prerequisite for their further use in clinical practice. Moreover, by using predefined search

procedures, we were able to isolate and compare the effect of several components of search strategies in relation to a specific clinical question. As we focused on retrieval, we considered each relevant article to be equally important, although this may not be the case. In particular, we did not examine selection bias attributable to suboptimal identification of relevant articles. Finally, our results apply only to therapeutic interventions, assessed by clinical trials, and cannot be generalized to questions on prevention or diagnosis [32]. For the latter, similar studies could be conducted using alternative Clinical Queries filters designed to improve the retrieval of high-quality studies on diagnosis or prognosis [26].

Among its strengths, our design allowed an unbiased comparison of strategies, by controlling important sources of variation such as clinicians' searching skills or previous knowledge [29,30]. This exploration was broader and more systematic than in previous studies. Moreover, the tested search strategies relied on search components that are easily applicable in practice, as they are directly implemented in PubMed. Finally, we used reputable systematic reviews to identify relevant studies on specific clinical topics.

Conclusion and Future Prospects

Countless PubMed searches are performed daily by clinicians, but the effectiveness of this activity is poorly understood [20]. We have shown that a well-formulated PIC query used with the Clinical Queries narrow filter was most likely to retrieve relevant clinical trials within readable outputs. These results can help clinicians build more effective strategies to answer their questions at the point of care, and thus reduce the gap between evidence from clinical trials and its actual implementation in practice. Further research should focus on the performance and clinical usefulness of selected strategies when they are performed in real practice [45]. Meanwhile, PubMed remains a perfectible tool. Areas of improvement include the development of content filters for specific clinical disciplines [46], or the implementation of new search interfaces that help clinicians formulate effective queries [24] and conduct parallel searches combining methodological and content filters at will.

Acknowledgments

All persons who contributed significantly to this project are coauthors.

There was no external funding

Authors' Contributions

Study concept, design: Agoritsas, Merglen, Perneger. *Acquisition of data:* Agoritsas, Merglen. *Statistical analysis:* Agoritsas, Courvoisier, Combescure. *Interpretation of data:* Agoritsas, Merglen, Courvoisier, Combescure, Garin, Perrier, Perneger. *Drafting of the article:* Agoritsas. *Critical revision of the article for important intellectual content:* Merglen, Courvoisier, Combescure, Garin, Perrier, Perneger. *Supervision:* Perneger.

Agoritsas had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed description of the 30 Cochrane systematic reviews on therapeutic interventions, used to identify relevant studies on 30 clinical questions.

[PDF File (Adobe PDF File), 85KB - [jmir_v14i3e85_app1.pdf](#)]

Multimedia Appendix 2

Two examples of search term extraction forms for one complex (review #1) and one simpler (review #7) clinical questions.

[PDF File (Adobe PDF File), 68KB - [jmir_v14i3e85_app2.pdf](#)]

Multimedia Appendix 3

Detailed search terms for 30 clinical questions, extracted from the abstract of the corresponding Cochrane systematic review according to the PICO framework.

[PDF File (Adobe PDF File), 69KB - [jmir_v14i3e85_app3.pdf](#)]

References

1. Straus SE, Richardson WS, Glasziou P, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM. 3rd edition. Edinburgh: Elsevier/Churchill Livingstone; 2005.
2. González-González AI, Dawes M, Sánchez-Mateos J, Riesgo-Fuertes R, Escortell-Mayor E, Sanz-Cuesta T, et al. Information needs and information-seeking behavior of primary care physicians. *Ann Fam Med* 2007;5(4):345-352 [FREE Full text] [doi: [10.1370/afm.681](#)] [Medline: [17664501](#)]
3. Green ML, Ciampi MA, Ellis PJ. Residents' medical information needs in clinic: are they being met? *Am J Med* 2000 Aug 15;109(3):218-223. [Medline: [10974185](#)]
4. Graber MA, Randles BD, Ely JW, Monnahan J. Answering clinical questions in the ED. *Am J Emerg Med* 2008 Feb;26(2):144-147. [doi: [10.1016/j.ajem.2007.03.031](#)] [Medline: [18272092](#)]
5. Straus SE, McAlister FA. Evidence-based medicine: a commentary on common criticisms. *CMAJ* 2000 Oct 3;163(7):837-841. [Medline: [11033714](#)]
6. Glasziou P, Burls A, Gilbert R. Evidence based medicine and the medical curriculum. *BMJ* 2008;337:a1253. [Medline: [18815165](#)]
7. US National Library of Medicine, National Institutes of Health. PubMed URL: <http://www.ncbi.nlm.nih.gov/sites/entrez?otool=ichbfmglib> [accessed 2011-11-30] [WebCite Cache ID 63a0b5P6K]
8. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010 Sep;7(9):e1000326 [FREE Full text] [doi: [10.1371/journal.pmed.1000326](#)] [Medline: [20877712](#)]
9. Haynes RB. Of studies, syntheses, synopses, summaries, and systems: the “5S” evolution of information services for evidence-based healthcare decisions. *Evid Based Med* 2006 Dec;11(6):162-164. [doi: [10.1136/ebm.11.6.162-a](#)] [Medline: [17213159](#)]
10. UpToDate, Inc. 2011. Helping Clinicians Provide the Best Patient Care URL: <http://www.uptodate.com/index> [accessed 2011-11-30] [WebCite Cache ID 63a0n7Fqj]
11. Thiele RH, Poiri NC, Scalzo DC, Nemergut EC. Speed, accuracy, and confidence in Google, Ovid, PubMed, and UpToDate: results of a randomised trial. *Postgrad Med J* 2010 Aug;86(1018):459-465. [doi: [10.1136/pgmj.2010.098053](#)] [Medline: [20709767](#)]
12. Hoogendam A, Stalenhoef AF, Robbé PF, Overbeke AJ. Answers to questions posed during daily patient care are more likely to be answered by UpToDate than PubMed. *J Med Internet Res* 2008;10(4):e29 [FREE Full text] [doi: [10.2196/jmir.1012](#)] [Medline: [18926978](#)]
13. Banzi R, Cinquini M, Liberati A, Moschetti I, Pecoraro V, Tagliabue L, et al. Speed of updating online evidence based point of care summaries: prospective cohort analysis. *BMJ* 2011;343:d5856 [FREE Full text] [Medline: [21948588](#)]
14. Haynes RB. ACP Journal Club: the best new evidence for patient care. *ACP J Club* 2008 May 20;148(3):2. [Medline: [18489065](#)]
15. Patel MR, Schardt CM, Sanders LL, Keitz SA. Randomized trial for answers to clinical questions: evaluating a pre-appraised versus a MEDLINE search protocol. *J Med Libr Assoc* 2006 Oct;94(4):382-387. [Medline: [17082828](#)]
16. Bonis PA, Pickens GT, Rind DM, Foster DA. Association of a clinical knowledge support system with improved patient safety, reduced complications and shorter length of stay among Medicare beneficiaries in acute care hospitals in the United States. *Int J Med Inform* 2008 Nov;77(11):745-753. [doi: [10.1016/j.ijmedinf.2008.04.002](#)] [Medline: [18565788](#)]
17. McKibbon KA, Fridsma DB. Effectiveness of clinician-selected electronic information resources for answering primary care physicians' information needs. *J Am Med Inform Assoc* 2006;13(6):653-659 [FREE Full text] [doi: [10.1197/jamia.M2087](#)] [Medline: [16929042](#)]

18. Schilling LM, Steiner JF, Lundahl K, Anderson RJ. Residents' patient-specific clinical questions: opportunities for evidence-based learning. *Acad Med* 2005 Jan;80(1):51-56. [Medline: [15618093](#)]
19. Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Weir MA, Haynes RB, et al. Searching for medical information online: a survey of Canadian nephrologists. *J Nephrol* 2011;24(6):723-732. [doi: [10.5301/JN.2011.6373](#)] [Medline: [21360475](#)]
20. 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](#)]
21. Hoogendam A, Stalenhoef AF, Robbé PF, Overbeke AJ. Analysis of queries sent to PubMed at the point of care: observation of search behaviour in a medical teaching hospital. *BMC Med Inform Decis Mak* 2008;8:42 [FREE Full text] [doi: [10.1186/1472-6947-8-42](#)] [Medline: [18816391](#)]
22. Webster AC, Cross NB, Mitchell R, Craig JC. How to get the most from the medical literature: searching the medical literature effectively. *Nephrology (Carlton)* 2010 Feb;15(1):12-19. [doi: [10.1111/j.1440-1797.2009.01263.x](#)] [Medline: [20377765](#)]
23. Doig GS, Simpson F. Efficient literature searching: a core skill for the practice of evidence-based medicine. *Intensive Care Med* 2003 Dec;29(12):2119-2127. [doi: [10.1007/s00134-003-1942-5](#)] [Medline: [12955188](#)]
24. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak* 2007;7:16 [FREE Full text] [doi: [10.1186/1472-6947-7-16](#)] [Medline: [17573961](#)]
25. US National Library of Medicine, National Institutes of Health. PubMed Limits URL: <http://www.ncbi.nlm.nih.gov/pubmed/limits> [accessed 2011-11-30] [WebCite Cache ID 63a11ZmDN]
26. US National Library of Medicine, National Institutes of Health. PubMed Clinical Queries URL: <http://www.ncbi.nlm.nih.gov/pubmed/clinical> [accessed 2011-11-30] [WebCite Cache ID 63a17P1uh]
27. Haynes RB, McKibbin KA, Wilczynski NL, Walter SD, Werre SR, Hedges Team. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005 May 21;330(7501):1179. [doi: [10.1136/bmj.38446.498542.8F](#)] [Medline: [15894554](#)]
28. US National Library of Medicine, National Institutes of Health. 2011 Nov 8. Abridged Index Medicus (AIM or "Core Clinical") Journal Titles URL: <http://www.nlm.nih.gov/bsd/aim.html> [accessed 2011-11-30] [WebCite Cache ID 63a1LosQL]
29. Stark R, Helenius IM, Schimming LM, Takahara N, Kronish I, Korenstein D. Real-time EBM: from bed board to keyboard and back. *J Gen Intern Med* 2007 Dec;22(12):1656-1660. [doi: [10.1007/s11606-007-0387-x](#)] [Medline: [17922170](#)]
30. Gruppen LD, Rana GK, Arndt TS. A controlled comparison study of the efficacy of training medical students in evidence-based medicine literature searching skills. *Acad Med* 2005 Oct;80(10):940-944. [Medline: [16186614](#)]
31. Hoogendam A, de Vries Robbé PF, Stalenhoef AF, Overbeke AJ. Evaluation of PubMed filters used for evidence-based searching: validation using relative recall. *J Med Libr Assoc* 2009 Jul;97(3):186-193. [doi: [10.3163/1536-5050.97.3.007](#)] [Medline: [19626144](#)]
32. Kastner M, Wilczynski NL, McKibbin AK, Garg AX, Haynes RB. Diagnostic test systematic reviews: bibliographic search filters ("Clinical Queries") for diagnostic accuracy studies perform well. *J Clin Epidemiol* 2009 Sep;62(9):974-981. [doi: [10.1016/j.jclinepi.2008.11.006](#)] [Medline: [19230607](#)]
33. Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Med Res Methodol* 2006;6:33 [FREE Full text] [doi: [10.1186/1471-2288-6-33](#)] [Medline: [16848895](#)]
34. Delaney A, Bagshaw SM, Ferland A, Laupland K, Manns B, Doig C. The quality of reports of critical care meta-analyses in the Cochrane Database of Systematic Reviews: an independent appraisal. *Crit Care Med* 2007 Feb;35(2):589-594. [doi: [10.1097/01.CCM.0000253394.15628.FD](#)] [Medline: [17205029](#)]
35. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *JAMA* 1998 Jul 15;280(3):278-280. [Medline: [9676681](#)]
36. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med* 2007 Mar 27;4(3):e78 [FREE Full text] [doi: [10.1371/journal.pmed.0040078](#)] [Medline: [17388659](#)]
37. The Cochrane Collaboration. John Wiley & Sons, Inc. 2012. The Cochrane Library: Advanced Search URL: <http://onlinelibrary.wiley.com/o/cochrane/mainSearch?action=displayForm> [accessed 2012-05-21] [WebCite Cache ID 67ppfvtsm]
38. Poole P, Black PN. Mucolytic agents for chronic bronchitis or chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2010(2):CD001287. [doi: [10.1002/14651858.CD001287.pub3](#)] [Medline: [20166060](#)]
39. McKibbin KA, Wilczynski NL, Haynes RB, Hedges Team. Retrieving randomized controlled trials from medline: a comparison of 38 published search filters. *Health Info Libr J* 2009 Sep;26(3):187-202. [doi: [10.1111/j.1471-1842.2008.00827.x](#)] [Medline: [19712211](#)]
40. Higgins JPT, Green SE. The Cochrane Collaboration. 2011 Mar. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0, Section 6.4.2, Structure of a Search Strategy URL: <http://www.cochrane-handbook.org/> [accessed 2012-05-16] [WebCite Cache ID 67i0XLLQh]
41. Lin J, Wilbur WJ. PubMed related articles: a probabilistic topic-based model for content similarity. *BMC Bioinformatics* 2007;8:423 [FREE Full text] [doi: [10.1186/1471-2105-8-423](#)] [Medline: [17971238](#)]

42. R Development Core Team. R Foundation for Statistical Computing. 2008. The R Project for Statistical Computing URL: <http://www.r-project.org/> [accessed 2011-11-30] [WebCite Cache ID 63a1mUGlw]
43. Hersh WR, Crabtree MK, Hickam DH, Sacherek L, Friedman CP, Tidmarsh P, et al. Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. *J Am Med Inform Assoc* 2002;9(3):283-293 [FREE Full text] [Medline: [11971889](#)]
44. Vakkari P, Huuskonen S. Search effort degrades search output but improves task outcome. *J Am Soc Inf Sci Technol* 2012;63(4):657-670. [doi: [10.1002/asi.21683](#)]
45. Shariff SZ, Cuerden MS, Haynes RB, McKibbin KA, Wilczynski NL, Iansavichus AV, et al. Evaluating the impact of MEDLINE filters on evidence retrieval: study protocol. *Implement Sci* 2010;5:58 [FREE Full text] [doi: [10.1186/1748-5908-5-58](#)] [Medline: [20646295](#)]
46. Garg AX, Iansavichus AV, Wilczynski NL, Kastner M, Baier LA, Shariff SZ, et al. Filtering Medline for a clinical discipline: diagnostic test assessment framework. *BMJ* 2009;339:b3435 [FREE Full text] [Medline: [19767336](#)]

Abbreviations

AIM: Abridged Index Medicus

COPD: chronic obstructive pulmonary disease

IQR: interquartile range

MeSH: Medical Subject Headings

PIC(O): population, interventions, comparison, (outcomes)

PPV: positive predictive value

Edited by G Eysenbach; submitted 04.12.11; peer-reviewed by M Sampson, S Shariff, K Robinson; comments to author 05.02.12; revised version received 11.03.12; accepted 13.04.12; published 12.06.12.

Please cite as:

Agoritsas T, Merglen A, Courvoisier DS, Combescure C, Garin N, Perrier A, Perneger TV

Sensitivity and Predictive Value of 15 PubMed Search Strategies to Answer Clinical Questions Rated Against Full Systematic Reviews *J Med Internet Res* 2012;14(3):e85

URL: <http://www.jmir.org/2012/3/e85/>

doi: [10.2196/jmir.2021](#)

PMID: [22693047](#)

©Thomas Agoritsas, Arnaud Merglen, Delphine S Courvoisier, Christophe Combescure, Nicolas Garin, Arnaud Perrier, Thomas V Perneger. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org/>), 12.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

What do Web-Use Skill Differences Imply for Online Health Information Searches?

Markus A Feufel^{1*}, PhD; S Frederica Stahl^{1*}, MA

Harding Center for Risk Literacy, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany

* all authors contributed equally

Corresponding Author:

Markus A Feufel, PhD

Harding Center for Risk Literacy

Center for Adaptive Behavior and Cognition

Max Planck Institute for Human Development

Lentzeallee 94

Berlin, 14195

Germany

Phone: 49 3082406 ext 292

Fax: 49 3082406394

Email: feufel@mpib-berlin.mpg.de

Abstract

Background: Online health information is of variable and often low scientific quality. In particular, elderly less-educated populations are said to struggle in accessing quality online information (digital divide). Little is known about (1) how their online behavior differs from that of younger, more-educated, and more-frequent Web users, and (2) how the older population may be supported in accessing good-quality online health information.

Objective: To specify the digital divide between skilled and less-skilled Web users, we assessed qualitative differences in technical skills, cognitive strategies, and attitudes toward online health information. Based on these findings, we identified educational and technological interventions to help Web users find and access good-quality online health information.

Methods: We asked 22 native German-speaking adults to search for health information online. The skilled cohort consisted of 10 participants who were younger than 30 years of age, had a higher level of education, and were more experienced using the Web than 12 participants in the less-skilled cohort, who were at least 50 years of age. We observed online health information searches to specify differences in technical skills and analyzed concurrent verbal protocols to identify health information seekers' cognitive strategies and attitudes.

Results: Our main findings relate to (1) attitudes: health information seekers in both cohorts doubted the quality of information retrieved online; among poorly skilled seekers, this was mainly because they doubted their skills to navigate vast amounts of information; once a website was accessed, quality concerns disappeared in both cohorts, (2) technical skills: skilled Web users effectively filtered information according to search intentions and data sources; less-skilled users were easily distracted by unrelated information, and (3) cognitive strategies: skilled Web users searched to inform themselves; less-skilled users searched to confirm their health-related opinions such as "vaccinations are harmful." Independent of Web-use skills, most participants stopped a search once they had found the first piece of evidence satisfying search intentions, rather than according to quality criteria.

Conclusions: Findings related to Web-use skills differences suggest two classes of interventions to facilitate access to good-quality online health information. Challenges related to findings (1) and (2) should be remedied by improving people's basic Web-use skills. In particular, Web users should be taught how to avoid information overload by generating specific search terms and to avoid low-quality information by requesting results from trusted websites only. Problems related to finding (3) may be remedied by visually labeling search engine results according to quality criteria.

(*J Med Internet Res* 2012;14(3):e87) doi:[10.2196/jmir.2051](https://doi.org/10.2196/jmir.2051)

KEYWORDS

Public access to information; computer literacy; information-seeking behavior; patient education handout; human factors; digital divide; information retrieval

Introduction

Unlike most information environments in human history, the Internet provides almost unlimited and constantly changing information sources of varying and often unknown quality. These properties pose a challenge to people navigating the Internet, in particular if information contents are highly specialized as in the medical domain. Analyses of online sources show that online health information is of variable and in most cases low scientific quality [1,2]. In addition, studies suggest that a digital divide limits older and less-educated people's Web-use skills and, ultimately, their access to online information [3]. Cognizant of these challenges, researchers and both governmental and nongovernmental institutions have attempted to objectively map the quality landscape of the Internet [4-6]. These initiatives provide site recommendations and transparency criteria that can be used to evaluate websites, including information about data sources, authorship, or last update. Although health information seekers may be aware of such criteria, only "few...notice and later [remember] from which websites they retrieved information or who stood behind the sites" (page 576 [7]). Instead, Web users rely on and trust the first few results provided by widely used search engines, such as Google, Bing, or Yahoo! [7,8].

Apart from a general consensus that people are poor at searching and evaluating online information [9], little is known about how user characteristics, such as Web-use skill differences, affect online behaviors and may inform interventions to facilitate retrieval of good-quality information [10,11]. Moreover, most research into online behavior is limited to mock-up sites [12], restricted search spaces [13,14], and quantifiable outcome measures such as search time, number of search terms entered, or clicks [15]. To complement experimental research and generate first hypotheses about how to help health information seekers find good-quality information, we identified qualitative differences in skilled versus less-skilled Web users' (1) attitudes toward online health information, (2) technical skills, and (3) cognitive strategies for search and evaluation of online health information.

Methods

Participants

Participants were native German speakers recruited from the participant pool at the Max Planck Institute for Human Development in Berlin, Germany. To identify meaningful differences in participants' Web-use skills, we used demographic factors that have been shown to facilitate access and use of online information, including younger age [16] and higher education [17]. In addition, we used a Web-oriented digital literacy survey measure as a validated proxy for people's actual Web-use skills [18,19]. Finally, we characterized cohorts in terms of online behavior using items from the General Social Survey (GSS) [10], which provided a quantitative measure of people's health information-seeking behavior, offline and online, as well as information about Web access and overall time spent online.

Cohort Characteristics

The older cohort consisted of 12 individuals (6 women) who were at least 50 years of age (mean 65, SD 4.3 years). Most members of this cohort (9/12, 75%) were secondary school (*Realschule*) graduates; 3 participants had a university degree. The younger cohort consisted of 10 individuals (5 women) who were all younger than 30 years of age (mean 23, SD 3.3 years) and on average more highly educated than the less-skilled cohort. Of the 10 participants in this group, 9 had graduated from high school (*Gymnasium*) and 5 were pursuing a university degree.

Concerning health information-seeking behavior, all participants reported having searched for health information at least once in the previous year. To do so, all but 3 of the older participants (19/22, 86%) stated that they had searched online. Apart from these similarities, members of the younger cohort reported spending almost 6 times more time online (mean 1022, SD 773 minutes/week) than their older counterparts (mean 175, SD 120 minutes/week). We also found strong cohort differences in participants' Web-oriented digital literacy scores. Whereas older participants obtained on average about half of 130 possible literacy points (mean 63, SD 22 points), members of the younger cohort obtained about three-quarters of the possible points (mean 98, SD 21 points). Thus, the demographic cohort differences of age and education level as well as the Web-related differences of time spent online and digital literacy scores spoke to a general advantage of the younger cohort in terms of accessing and using online information [16,17]. We therefore refer to these cohorts as skilled and less-skilled Web users.

Materials

Most search tasks used to study online health information-seeking behavior have asked participants to search for facts, as in "What is the definition of being overweight?" [7,8]. However, prior research assumed [13,20,21] and "a great many health seekers say the resources they find on the Web have a direct effect on the decisions they make...and on their interactions with doctors" (page 3 [22]). Thus, we provided search scenarios that asked participants to search for facts, *as well as* inference scenarios that asked them to make decisions or recommendations based on information retrieved online (see Table 1).

We designed 3 search scenarios to capture the most common health facts searched online [23]: information about diseases and symptoms (searched by 66% of US adult Web users), information about prescription drugs (45%), and information about medical tests (16%). Similarly, we developed 4 inference scenarios capturing common motivations for inference-related online searches [24]. One inference scenario related to pregnancy (searched by 19% of US adult Web users) and 2 scenarios concerned vaccination decisions (16%). We added a self-treatment and diagnosis scenario because participants frequently commented on the phenomenon during the first 8 sessions and "there is great concern in the medical establishment that e-patients are self-diagnosing and self-medicating" (page 15 [23]). Surveys suggest that 18% of US adults use the Internet to self-diagnose [23].

Table 1. Search scenarios presented to participants.

Scenarios	Common search motivations identified in the literature ^a
Search scenarios	
A physician talked to you/your husband/your father about taking a PSA ^b test. The doctor explained that the validity of the test has been questioned by a number of physicians. How do you assess the validity of the test?	16% search for interpretations of medical test results [20]
Your doctor prescribes eszopiclone and lets you know that you might develop a bad aftertaste when taking the medication. You do not develop a bad aftertaste, but feel occasionally nauseous. Could nausea be a result of taking eszopiclone?	45% search for information on prescription drugs [20]
Your partner/parent has been told by his/her physician that he/she has an increased risk of a stroke. How would you recognize a stroke?	66% search for a particular illness or condition [20]
Inference scenarios	
Your daughter's/your sister's gynecologist recommends that she should get the gardasil vaccination. What would you advise her to do?	16% search for immunization- and vaccine-related information [21]
Close friends of yours are contacted by their pediatrician, who recommends that their child get an MMR ^c vaccination. Your friends discuss the matter with you. Should their child be vaccinated?	16% search for immunization- and vaccine-related information [21]
Your pregnant daughter's/friend's gynecologist suggests that she might want to undergo an amniocentesis. What would you advise her?	19% search for pregnancy-related information [21]
You wake up one morning with a swollen elbow. What could that be? How would you go about treating it?	18% state using the Internet to self-diagnose [20]

^a Search motivations are based on survey results provided by the Pew Internet & American Life Project. Percentages refer to representative samples of US Web users.

^b Prostate-specific antigen.

^c Measles-mumps-rubella.

Given that most of the scenarios target particular populations, we asked participants to search for information for friends or family members belonging to these populations. Research suggests that 81% of adult Web users search online information for others [23], so this change of perspective should not have reduced the perceived relevance of the scenarios. All study materials including consent forms were approved by the ethics committee at the Max Planck Institute for Human Development.

Data Collection and Procedure

Sessions lasted about 90 minutes and were conducted by one of the authors (FS) who observed participants throughout the sessions. To be as unobtrusive as possible, we explained the think-aloud method [25] to participants at the beginning of the sessions, asking them to verbalize thoughts during online searches. In case participants discontinued thinking aloud, a list of probing questions was devised. Examples include “How did you come up with your search term?”, “Why did you click on this link?”, and “What are you looking for on this site?” to probe for search implementation, site selection, and site navigation strategies, respectively. To record verbal protocols and on-screen behaviors, we used a MacPro laptop (Apple Inc, Cupertino, CA, USA) with iShowU screen-capture software (shinywhitebox ltd, Wellington, New Zealand).

After providing informed consent, participants were given one of the scenarios chosen at random. Without specifically requesting an Internet search, we asked them to report on how they would approach the scenario. This part was used to identify attitudes toward online health information and typical health information search patterns. If online search was not part of the

strategy they mentioned, they were queried at the end of the session about the role they assign to online health information. After this part, participants were shown the computer and how to open a blank browser window. Then, they were given the remaining 6 scenarios, consecutively and in random order. Participants were unrestricted in terms of search time, how to approach searches, and which websites to visit. After having worked on the scenarios, participants filled out the surveys used to characterize cohorts.

Data Analysis

Verbal protocols were transcribed immediately after each session. Both authors then independently coded protocols and screen-captured online behaviors. Diverging codes were discussed until consensus was reached. Codes reflected *attitudes* toward online health information, *technical skills* (based on observable on-screen behaviors), and *cognitive strategies* (based on verbal protocols). Attitudes were coded with respect to the *categories of information* participants would or would not search for online and their *rationales* for doing or not doing so. Technical skills were subdivided into *search implementation*, which refers to the website(s) accessed to begin the search and the generation of search term(s); *site selection*, which refers to how people decided which of the search results to follow up on in detail; and *site navigation*, which describes how participants traveled through a selected site. Following decision-making research [26], we subdivided cognitive strategies into *search intentions* used to structure information retrieval; *rules for information evaluation*, referring to how participants interpreted online information and modified their search; *stopping rules*,

defined by how participants chose to end their search; and *inference rules*, defined by how they integrated the retrieved information into a decision or recommendation.

Results

In the following, we describe the impact of Web-use skill differences on (1) user attitudes, (2) technical skills, and (3) cognitive search and evaluation strategies during online health information searches. Table 2 summarizes our findings, which we discuss in turn.

Table 2. Summary of results.

Outcome	General findings	Skill differences	
		Less-skilled cohort	More-skilled cohort
Attitudes			
	Participants hesitated to use online health data (20/22, 91%), in particular to make inferences	Rationale: the plethora of data cannot be managed (6/12, 50%)	Rationale: the quality of online data is low (8/10, 80%)
	Although data quality was a matter of concern...	...keywords may override distrust toward sites (6/12, 50%)	...once accessed, data from any site were used (4/10, 40%)
Technical skills			
Search implementation	Reliance on search engine Google.de (22/22, 100%)	No differences observed	Few entered URLs directly (2/10, 20%)
	Use of search terms (20/22, 91%); often misspelled; rarely corrected	83% used a single search term (10/12); 67% used search term suggestions (8/12); 17% used natural language phrases (2/12)	100% used two or more search terms (10/10)
Site selection	Reliance on first 5 links on first search engine result page (20/22, 91%)	Selection based on keywords (often unrelated to original search) inferred from links/excerpts (7/12, 58%)	Selection based on data sources inferred from URL or links/excerpts (9/10, 90%)
Site navigation	Relevance of website contents was appraised	Text was read rather than scanned. In 23% of searches (14/61) links were followed up.	Text was scanned for keywords. In 6% of searches (3/49), links were followed up.
	Information was rarely cross-referenced	No cross-referencing or use of tabs (0/12, 0%).	Use of multiple tabs to compare results (2/10, 20%).
Cognitive strategies (based on inference scenarios only)			
Search intentions	People searched for online contents related to personal, a priori opinions, knowledge, cues, or expert opinions.	Distribution of intentions: 70% a priori opinions (21/30); 10% cues (3/30); 10% knowledge (3/30); 10% expert opinions (3/30)	Distribution of intentions: 14% a priori opinions (4/28); 29% cues (8/28); 29% knowledge (8/28); 29% expert opinions (8/28)
Information evaluation	Information was trusted if consistent with search intentions—that is, if...	...a website confirmed a priori opinions (21/30, 70%) or yielded search contents (9/30, 30%)	...a website confirmed a priori opinions (4/28, 14%) or yielded search contents (24/28, 86%).
Stopping rule	Search was stopped once the first piece of online information satisfied search intentions	No participant further cross-referenced (0/12, 0%)	20% further cross-referenced information (2/10)
Inference rule	Participants were hesitant to make inferences based on online searches, except when they searched to confirm personal, a priori opinions	In 73% of the inference queries, inferences were made based on a priori opinions (22/30). In 27% no inferences were made (8/30).	In 14% of the inference queries, inferences were made based on a priori opinions (4/28); in 7% based on cues (2/28). In 79% no inferences were made (22/28).

Attitudes Toward Online Health Information Seeking

Participants identified clear boundaries as to when and for what purpose they would access online health information. For instance, all participants who were asked to self-diagnose and treat a swollen elbow (14/22) made statements such as “I don’t trust the Internet when it comes to symptoms...home remedies maybe, but diagnoses definitely not.” In fact, 7 of the 14

participants asked to self-diagnose (50%) refused to search online; the other half searched for “swollen elbow” but, after having scanned the first results, decided that looking up symptoms online was inappropriate and discontinued their search.

Similarly, during the first scenario, when not explicitly asked to use the Internet, 7 of the 12 less-skilled participants (58%)

did not consider the Web; only 1 (8%) referred to it as the primary source of health information. Although all 10 skilled participants mentioned the Internet, only half identified it as their primary source. In general, 20 of 22 participants (91%) made statements reflecting distrust toward online sources: “First I would turn to my doctor; then to family members, maybe close friends. A distant third would be online research....a very distant third” or “My trust in the Internet is negatively correlated to the importance I ascribe to an issue: the more important the issue, the less I trust, and thus rely on, the Internet.”

When we asked participants to actually surf the Web (after completing the first scenario), verbal protocols suggested different reasons for why more- versus less-skilled Web users were hesitant to access online health information. Half of the less-skilled users (6/12) made statements such as “Online one gets easily carried away” and “I shouldn’t search online by myself or I will be bombarded by all this information. One has to be a professional to ask the right questions.” Thus, less-skilled participants expressed concern about the amount of available information and their lack of navigation skills. In contrast, 8 of the 10 skilled participants (80%) made statements such as “It’s difficult to find valid health information online...anyone can put a site online.” A total of 5 participants (50%) stated that health was too important a topic to rely solely on online resources: “You can’t always trust Google. I’d rather ask a doctor or pharmacist.” All 10 skilled Web users (100%) questioned the quality of and trust they had in online sources at least once during their online searches.

The hesitant attitude toward online health information seems at odds with earlier studies, which assumed that online information influences medical decision making [13,20-22]. This is only an apparent contradiction, however. Once participants accessed a website in our study, their concerns about data quality vanished. Among the 12 less-skilled users, only 2 (17%) mentioned quality issues explicitly once online, and 6 of them (50%) doubted a source but, after identifying an interesting keyword, clicked on the link and reasoned about the content of the site nonetheless. Similarly, skilled participants only infrequently referred to trust once they accessed a page. In 4 of their 20 searches, 4 participants read and considered information in their reasoning, even though they had expressed distrust of the source before deciding to visit the website. Thus, although people voice concerns about data quality issues, once they access a website, even skilled Web users are preoccupied with processing website contents.

Technical Skills Based on Observed Online Behaviors

Overall, participants performed 110 online searches: 52 (47%) were related to search scenarios and 58 (53%) were related to inference scenarios (see Table 3). Prior observational research has found that people may search from 5 to 20 minutes before they stop [7,8]. Our results suggest search times of about 5 minutes for search scenarios and 6 minutes for inference scenarios. In general, more-skilled participants spent about 2 minutes less time searching than their less-skilled counterparts.

Table 3. Number of online searches and mean search time by search scenario and cohort type.^a

Scenario type	Number of online searches			Mean search time (minutes:seconds)		
	Less-skilled cohort	More-skilled cohort	Total	Less-skilled cohort	More-skilled cohort	Total
Search scenarios	31	21	52	5:04	4:17	5:16
Validity of PSA ^b test	10	9	19	6:32	4:31	5:48
Drug side effects	11	9	20	6:25	4:37	5:31
Stroke symptoms	10	3	13	3:50	2:34	3:35
Inference scenarios	30	28	58	7:53	4:57	6:22
MMR ^c vaccination	12	9	21	9:03	6:00	7:45
Gardasil vaccination	8	9	17	8:21	5:13	6:29
Amniocentesis	10	10	20	6:04	3:46	4:49
Self-diagnosis ^d	(3)	(4)	(7)	(2:15)	(2:34)	(2:27)
Overall	61	49	110	6:45	4:40	5:47

^a Given that the first scenario was not followed up with actual online searches, there is a total of $(6 - 1) \times 22 = 110$ searches.

^b Prostate-specific antigen.

^c Measles-mumps-rubella.

^d In response to the self-diagnosis scenario, 7 searches were started but not completed, which are added in parentheses but not in the totals.

Search Implementation

In line with prior observational work [7,8], participants mainly relied on search engines to access online information (in 107/110, 97% of the searches). All 22 participants showed a

strong preference for the search engine Google.de. Also echoing earlier findings, 91% (20/22) of the participants searched for health information using single-word search terms, which were often left uncorrected if misspelled, even when the search engine provided alternative, correct spellings. This cohort comparison

extends prior research by revealing differences with respect to the kinds and specificity of the search terms generated.

Of the 12 less-skilled Web users, 10 (83%) entered single-word search terms, often the disease or medication discussed in the scenario. The remaining 2 (17%) participants entered full natural language phrases, such as “How would I recognize a stroke?” Consequently, the majority of the less-skilled participants received rather general search results. Of the 12 less-skilled participants, 8 (67%) considered the search term options suggested by Google’s search bar to help specify search terms. They did not understand, however, how that would change their search results. For example, 1 participant searched for *eszopiclone* (a sleeping drug). When Google suggested the search term *side effects* to specify the query, the participant accepted the suggestion but, looking at the results, exclaimed “Oh, look at all those side effects. Every title listed includes side effects as a catchphrase...No, I would definitely talk to my doctor [before taking this medication].” In contrast, all 10 skilled Web users performed Google searches based on two or more single-word search terms. They entered a disease or medication and specified the category of information they were looking for, such as “eszopiclone side effects”, “HPV vaccinations MD opinions”, or “MMR vaccine tolerance.” They mostly ignored search terms suggested by the search engine, indicating prespecified search intentions.

Site Selection

About 91% (20/22) of all participants chose sites ranked among the first 5 links on the first result page. Participants left the first page provided by the search engine in only 7 of 107 (7%) Google searches. All of these 7 searches were related to the drug *eszopiclone*, a type of sleeping pill, available in the United States but not in Germany, so that the websites provided on Google.de were all in English. Thus, all of the 7 searchers leaving the first result page did so in hope of finding a German-language page.

Apart from a focus on the first result page, cohorts selected links differently. More than half of the less-skilled Web users (7/12, 58%) scanned link titles and excerpts provided by Google for keywords that triggered their interest. The remaining 5 (42%) participants occasionally ignored search engine titles and excerpts and identified keywords on websites. The keywords participants identified were often unrelated to their original search term entry. All 12 participants entered a disease as the search term at least once, read about side effects in an excerpt or title, and then decided to visit that site, without having actively searched for side effects.

Of the 10 skilled Web users, 9 (90%) decided which site to follow up by information source rather than keywords. They either directly recognized the source in the case of commonly known websites such as Netdoktor.de or extrapolated the type of source (eg, patient reports, health forums) from link titles, excerpts, or URLs. All members of this cohort relied at least once on Wikipedia.de to obtain a general overview; 6 (60%) of them expressed distrust toward available health forums and frequently chose not to visit a forum, even if it was ranked among the first entries on the results page.

Site Navigation

Once they selected a website, all participants started scanning the text to confirm its relatedness to the search topic. Cohorts differed in their scanning strategies. Less-skilled Web users started reading the text and followed the links provided on the websites. In 14 of 61 searches (23%), less-skilled participants accessed more than one page of a chosen website or looked at the table of contents. In addition, they considered information that was unrelated to their initial queries. For example, to find evidence concerning amniocenteses for a pregnant friend, 5 of the 12 (42%) less-skilled participants clicked on Netdoktor.de. On this website, the sidebars listed infomercials about ultrasound tests for pregnant women. Although it was not listed as an alternative to amniocentesis, the 5 participants followed the link and eventually suggested that their friend get an ultrasound as well.

Skilled participants used a more focused strategy to navigate websites and satisfy search intentions. In only 3 of 49 (6%) searches did skilled Web users access more than the first page of a chosen website. In all 3 instances, the original website was Wikipedia. Of the 10 skilled users, 4 (40%) commented that Wikipedia provided a good overview of any given topic and was a good place to start a search. To avoid reading too much text, 4 of the 10 (40%) skilled participants used the shortcut CTRL+F to bring up the Find utility and indentify search-related keywords on the website. To compare website contents, 2 (20%) of the skilled participants used multiple tabs to shuttle between websites.

Cognitive Strategies Based on Verbal Protocols

Search Intentions

There were marked differences in how the cohorts approached inference scenarios (see Table 4). In 21 of 30 (70%) instances, less-skilled participants based their search on a priori *opinions* toward the diseases, medications, or procedures mentioned in the scenarios, such as “I don’t like medications so I wouldn’t take it” or “I’m very skeptical about vaccines.” In the extreme, a participant self-identifying as an anthroposophist exclusively looked up websites about his philosophy of health. This tendency was so strong that for 6 of 30 inference-related queries (20%), less-skilled participants expressed a preference not to search online at all (and started searching only after having been prompted by the researcher) because they had already made up their mind about the scenario. Less-skilled Web users rarely searched for biomedical knowledge (in 3 of 30 queries), decision-relevant cues (3 of 30), or expert opinions (3 of 30).

More-skilled participants, on the other hand, searched to confirm a priori opinions in only 4 of 28 (14%) inference-related queries. In the 3 inference scenarios, members of this cohort generally used a more varied set of search strategies than their less-skilled counterparts (2.1 versus 1.3 different strategies) and more often searched for biomedical *knowledge* of diseases, medications, or procedures (in 8 of 28, 28% of the queries; eg, “Let’s see how this medication works.”); decision-relevant *cues* (8 of 28 queries; eg, “Well, I guess the question is what are the side effects?”); or *expert opinions* by health professionals, governments, or health organizations (8 of 28 queries; eg, “What

kind of doctor performs a PSA test? Urologist, right? Let's look what they say.""). These search intentions echo the search implementation technique observed in this more-skilled cohort. Their search terms first specified the medical condition or procedure and then an information category depending on how

they intended to approach the scenario by adding one or more search terms related to biomedical knowledge (eg, "physiology"), inference-relevant cues (eg, "side effects"), or "MD opinions".

Table 4. Participants' type of online search and evaluation strategy for inference scenarios.^a

Inference scenario	Strategy type ^b												Average
	1	2	3	4	5	6	7	8	9	10	11	12	
Less-skilled participants													
MMR ^c vaccine	K	O	O	O	O	O	K	O	O	O	O	EO	
Gardasil vaccine	O	O	O			O	C	O		O		EO	
Amniocentesis	K		O	O	O		C	O	O	O	C	EO	
No. of strategy types	2	1	1	1	1	1	2	1	1	1	2	1	1.3
More-skilled participants													
MMR vaccine	EO	K	EO	C	K		EO	EO	K	K			
Gardasil vaccine	C	C	EO	O	K	EO	C		O	K			
Amniocentesis	O	C	K	C	K	C	EO	O	EO	C			
No. of strategy types	3	2	2	2	1	2	2	2	3	2			2.1

^a Strategies in verbal protocols that were not followed up with actual online searches (ie, in response to the first and the self-diagnosis scenarios) are not included in this table.

^b O = participants searched to confirm *a priori* opinions, K = participants searched for biomedical knowledge, C = participants searched for decision-relevant cues, EO = participants searched for expert opinions on a topic.

^c Measles-mumps-rubella.

Information Evaluation

Confirming survey studies, we found that consistency encouraged participants to consider and trust Web content [23]. However, the kind of consistency participants were looking for was not between websites, as previously suggested, but between information retrieved online and search intentions. In 21 of 30 (70%) inference-related queries, less-skilled users evaluated information by relating it to their opinions about health issues. They trusted websites when they found that "my opinion is reflected here" or continued to search because they believed "there must be people out there who don't believe in vaccines either." Conflicting information was ascribed to bias on the part of the author(s): "Of course doctors support vaccinations. They make money off them." Alternatively, participants downplayed information disagreeing with their opinions. For instance, on visiting a site about the side effects of vaccinations, a participant who was generally pro vaccine stated: "I know there is a discussion about this procedure, but life is full of risks."

Skilled participants searched and evaluated information that concurred with their more varied repertoire of search intentions or strategies. For example, when two sites listed contradicting side effects, 1 participant searched for biomedical explanations to better understand the side effects in question. In a similar instance, another participant changed the cue he considered relevant. On finding two sites describing the same side effect as either infrequent or frequent, the participant stopped evaluating the procedure based on side effects and aimed at identifying whether he really needed the procedure.

Interestingly, none of the skilled participants who searched for expert opinions (see Table 4) acknowledged the possibility of conflicting opinions, nor did they continue searching after retrieving a first opinion.

Stopping Rule

The stopping rule participants in both cohorts commonly used was closely linked to their search intentions. That is, once they found the first piece of information that satisfied their search intentions (in the case of more-skilled participants) or was consistent with their health-related opinions (in the case of less-skilled participants), participants stopped searching. Only 2 of the more-skilled participants (20%) further cross-referenced such information once they had retrieved it.

Inference Rule

As stated above, participants were generally hesitant to rely on the Internet to make decisions or recommendations. However, less-skilled Web users based 73% (22/30) of their recommendations or decisions on *a priori* opinions, which is reflected in statements such as "Yes, I would get vaccinated; I'm generally in favor of vaccinations." Among the 10 more-skilled participants, 2 women expressed *a priori* opinions about the reproduction-related scenario and 2 men were generally in favor of vaccinations. In these 4 out of 28 (14%) instances, participants made recommendations based on their opinions. More-skilled Web users made recommendations based on inference-relevant cues in only 2 of 28 (7%) instances. For example, 1 participant referred to rare side effects as a reason to get vaccinated: "According to Wikipedia I would get

vaccinated...the side effects are not that bad after all.” In all other instances, participants reasoned about the online information but did not arrive at a conscious plan of action or recommendation.

Discussion

This study described the impact of Web-use skill differences on (1) user attitudes, (2) technical skills, and (3) cognitive search and evaluation strategies during online health information searches. After summarizing the results, we will present first hypotheses concerning interventions that, based on our findings, may facilitate the search for good-quality online health information.

First, both more- and less-skilled Web users were hesitant to use the Internet when making medical decisions, although for different reasons. Whereas poorly skilled Web users were concerned about managing data quantity, more-skilled users were concerned about its quality. However, once participants accessed online information, concerns about data quality vanished, independent of skill level. This suggests that, independent of actual Web-use skills, support interventions should focus on Web users' search implementation and site selection efforts.

Second, in terms of technical skills, skilled Web users effectively implemented and filtered information according to search intentions and data sources (eg, patient blog versus health forum). Less-skilled users had difficulties translating search intentions into search terms. Although they intended to search for relevant keywords on search engine result pages, less-skilled users were easily distracted by interesting but query-unrelated information, therefore often depending on coincidental findings. Again, this suggests that interventions should focus on search implementation and site selection efforts, in particular for users with poor Web-use skills.

Third, both cohorts evaluated search results based on consistency, although in different ways. Whereas most skilled Web users stopped a search once they found the first website providing the contents they were looking for, most less-skilled users posed less-specific queries and trusted information that was consistent with their health beliefs. Neither cohort systematically identified the quality of the retrieved information. This confirms studies suggesting that Web-use skills are necessary but not sufficient to guarantee access to and use of good-quality health information [27].

Limitations

The current study has several limitations. First, given that our aim was to account for qualitative differences in online health information seekers' attitudes, technical skills, and cognitive strategies, we were able to observe only a small number of individuals, which, in turn, limited our ability to identify meaningful differences at a more fine-grained level. Thus, as an observational and qualitative study, our work should be followed up with controlled experiments and larger sample sizes. Second, the degree of distrust toward online information may be specific to German health information seekers given that Germany has a universal health care system, and visiting and asking his or her doctor questions usually does not incur any costs for the patient. Future studies should corroborate our findings related to attitudes toward online information in different cultures and health care systems. Third, we used demographic and Internet-specific variables to identify reliable differences in Web-use skill. However, age and educational differences between cohorts also imply differences in prior exposure to and understanding of health issues, which may have equally affected the observed distribution of search and evaluation strategies across cohorts (see Table 4). Future studies should clarify whether older Web users mainly search to confirm a priori opinions based on richer health experiences or to compensate for poor Web-use skills. Fourth, although we carefully designed the scenarios to match common search motivations (see Table 1), they may have seemed irrelevant to some of the participants. Thus, our findings may differ from online behavior motivated by real health concerns. Also, the researcher conducting study sessions was careful to let participants know they had all the time they wanted to search. However, we cannot preclude that her presence might have affected participants' online behavior.

Implications

Based on the findings, we summarized avenues for interventions in Table 5, which we elaborate in turn. To prevent inequalities in terms of data access, interventions should focus on improving people's basic Web-use skills. A particular focus of these interventions should be skills related to search implementation and site selection efforts. To help people identify good-quality information independent of Web-use skills and simple stopping rules, search engine results should be visually labeled according to quality criteria.

Table 5. Avenues for interventions based on findings related to Web-use skill differences.

Finding	Intervention
Less-skilled health information seekers are concerned with managing data quantity.	Specify search terms and information categories to be searched for or use natural language phrases.
Less-skilled health information seekers pay little attention to information source when selecting websites.	Restrict search results to trusted sites by adding the command “site:” to queries.
Most health information seekers stop searching after finding a first piece of evidence satisfying search intentions, without cross-referencing.	Use multiple tabs in a browsing window to facilitate comparison of search engine results.
Web-use skills are not sufficient to guarantee access to and use of good-quality health information [27].	Visually label search engine results according to quality criteria.

Specify Search Words

Less-skilled users stated information overload as a major obstacle to using the Internet. Skilled Web users concatenated two or more search terms to specify the topic (eg, a medication) and the information category (eg, side effects) they were looking for. Specifying search intentions helped them constrain search results and avoid unrelated information. A similar effect was achieved by 2 less-skilled participants who entered whole natural language phrases instead of search terms (eg, “What are the side effects of eszopiclone?”). This intuitive strategy may be appropriate for users with low Web-use skills, helping them translate search intentions into targeted queries.

Restrict Search to Trusted Sites

Less-skilled participants paid little attention to information source when selecting websites. To avoid low-quality evidence, users may add the “site:” command to search terms (eg, “search term site: Netdoktor,” “search term site: NIH”), which yields results stemming exclusively from the identified website, circumventing the need to identify trustworthy sites from a plethora of options.

Use Multiple Tabs to Compare Findings

Neither cohort systematically cross-referenced online information. However, 2 of the skilled participants’ navigation strategies may provide a remedy. Moving from sequential to parallel browsing, they opened several websites they considered relevant, each in a new tab, and then compared contents. The

parallel browsing function may facilitate cross-referencing behavior.

Visually Label Search Engine Results

Both cohorts used simple search and stopping rules. They consulted only a few links and stopped searching once they retrieved the first piece of information that was consistent with their search intentions. Thus, one way to facilitate access to good-quality information independent of Web-use skill level is to visually structure result rankings according to quality criteria. Although search engines may use quality criteria to rank order search results, it is currently difficult to assess which of the displayed search results link to good-quality information. In 2010, Google USA partnered with the US National Institutes of Health (NIH) to change this situation [28]. When users of Google.com search diseases or selected medications, the first and visually separated result links to an NIH-sponsored website. Alas, similar ventures do not yet exist for search algorithms in other countries. In Germany, for example, searching for brand name drugs on Google yields links to pharmaceutical companies, industry-sponsored websites, or advertising. A visual acknowledgment of the quality structure underlying search results assures that the first information that health information seekers come across and select is from high-quality sources. Given that results delivered by branded search engines are generally trusted [8], this increases the chances for Web users to find good-quality online information, independent of Web-use skills or the use of simple cognitive strategies.

Acknowledgments

Funding: Funding was provided by the Harding Center for Risk Literacy at the Max Planck Institute for Human Development, Berlin, Germany. Researchers were scientifically independent from the funding organization.

Ethics approval: This study, all materials, and the consent form were approved by the ethics committee at the Max Planck Institute for Human Development, Berlin, Germany.

Authors' Contributions

Both authors contributed equally to the conception and design of the study, analysis and interpretation of data, drafting and critical revision of the article, and final approval of the version to be published. As the guarantor Markus A Feufel accepts full responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish.

Conflicts of Interest

None declared.

References

1. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients to make sense of health statistics. *Psychol Sci Public Interest* 2007;8(2):53-96 [FREE Full text] [doi: [10.1111/j.1539-6053.2008.00033.x](https://doi.org/10.1111/j.1539-6053.2008.00033.x)]
2. Neumeyer-Gromen A, Bodemer N, Müller S, Gigerenzer G. Ermöglichen Medienberichte und Broschüren informierte Entscheidungen zur Gebärmutterhalskrebsprävention. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 2010;54:1197-1210. [doi: [10.1007/s00103-011-1347-5](https://doi.org/10.1007/s00103-011-1347-5)]
3. Hargittai E. Second-level digital divide: differences in peoples' online skills. *First Monday* 2002 Apr;7(4) [FREE Full text]
4. Eysenbach G. Infodemiology: The epidemiology of (mis)information. *Am J Med* 2002 Dec 15;113(9):763-765. [Medline: [12517369](https://pubmed.ncbi.nlm.nih.gov/12517369/)]
5. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 2002;287(20):2691-2700. [Medline: [12020305](https://pubmed.ncbi.nlm.nih.gov/12020305/)]
6. Gagliardi A, Jadad AR. Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination. *BMJ* 2002 Mar 9;324(7337):569-573. [Medline: [11884320](https://pubmed.ncbi.nlm.nih.gov/11884320/)]

7. 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. [Medline: [11884321](#)]
8. Hargittai E, Fullerton F, Menchen-Trevino E, Thomas K. Trust online: young adults' evaluation of Web content. *Int J Commun* 2010;4:468-494 [[FREE Full text](#)]
9. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. *Cyberpsychol Behav* 2004 Oct;7(5):497-510. [Medline: [15667044](#)]
10. Hargittai E. Beyond logs and surveys: in-depth measures of people's online skills. *J Am Soc Inf Sci* 2002;53:1239-1244. [doi: [10.1002/asi.10166](#)]
11. LaRue EM. *BMJ*. 2003. An expertise not utilized URL: <http://www.bmj.com/content/324/7337/573?tab=responses> [accessed 2012-01-05] [[WebCite Cache ID 64THQy2rP](#)]
12. Flanagin AJ, Metzger MJ. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media Soc* 2007;9:319-342. [doi: [10.1177/1461444807075015](#)]
13. Betsch C, Renkewitz F, Betsch T, Ulshöfer C. The influence of vaccine-critical websites on perceiving vaccination risks. *J Health Psychol* 2010 Apr;15(3):446-455. [doi: [10.1177/1359105309353647](#)] [Medline: [20348365](#)]
14. Kienhues D, Stadler M, Bromme R. Dealing with conflicting or consistent medical information on the Web: when expert information breeds laypersons' doubts about experts. *Learning Instr* 2011;21:193-204. [doi: [10.1016/j.learninstruc.2010.02.004](#)]
15. Buckingham Shum S, McKnight C. World Wide Web usability: introduction to this special issue. *Int J Hum Comput Stud* 1997;47:1-4. [doi: [10.1006/ijhc.1997.0132](#)]
16. Freese J, Rivas S, Hargittai E. Cognitive ability and Internet use among older adults. *Poetics* 2006;34:236-249. [doi: [10.1016/j.poetic.2006.05.008](#)]
17. Hargittai E, Hinnant A. Digital inequality: differences in young adults' use of the Internet. *Commun Res* 2008;35:602-621. [doi: [10.1177/0093650208321782](#)]
18. Hargittai E. Survey measures of web-oriented digital literacy. *Soc Sci Comput Rev* 2005;23:371-379. [doi: [10.1177/0894439305275911](#)]
19. Hargittai E. An update on survey measures of web-oriented digital literacy. *Soc Sci Comput Rev* 2008;27:130-137. [doi: [10.1177/0894439308318213](#)]
20. Anderson JG, Rainey MR, Eysenbach G. The impact of CyberHealthcare on the physician-patient relationship. *J Med Syst* 2003 Feb;27(1):67-84. [Medline: [12617199](#)]
21. Eysenbach G. The impact of the Internet on cancer outcomes. *CA Cancer J Clin* 2003;53(6):356-371. [Medline: [15224975](#)]
22. Fox S, Rainie L. The online health care revolution. Washington, DC: Pew Internet & American Life Project; 2000 Nov 26. URL: http://www.pewinternet.org/~media/Files/Reports/2000/PIP_Health_Report.pdf.pdf [accessed 2012-06-06] [[WebCite Cache ID 68Din3nts](#)]
23. Fox S, Rainie L. Vital decisions. Washington, DC: Pew Internet & American Life Project; 2002 May 22. URL: http://www.pewinternet.org/~media/Files/Reports/2002/PIP_Vital_Decisions_May2002.pdf.pdf [accessed 2012-06-06] [[WebCite Cache ID 68DionsJB](#)]
24. Fox S. Health topics: 80% of Internet users look for health information online. Washington, DC: Pew Internet & American Life Project; 2011 Feb 01. URL: http://pewinternet.org/~media/Files/Reports/2011/PIP_HealthTopics.pdf [accessed 2012-06-06] [[WebCite Cache ID 68Ditvt09](#)]
25. Ericsson KA, Simon HA. Verbal reports as data. *Psychol Rev* 1980;87:215-251. [doi: [10.1037/0033-295X.87.3.215](#)]
26. Gigerenzer G, Todd PM, ABC Research Group. Simple Heuristics That Make us Smart. New York, NY: Oxford University Press; 2001.
27. Hargittai E. Digital na(t)ives? Variation in Internet skills and uses among members of the "net generation". *Sociol Inq* 2010;80(1):92-113 [[FREE Full text](#)]
28. Andrews J. Pharma Marketing Insight. 2010 Aug. Google Partners With NIH: Government Gets Top Spot URL: http://www.contactdd.com/DDDocuments/August_2010.pdf [accessed 2012-01-05] [[WebCite Cache ID 64T8x1eKC](#)]

Abbreviations

MMR: measles-mumps-rubella
NIH: National Institutes of Health
PSA: prostate-specific antigen

Edited by G Eysenbach; submitted 11.01.12; peer-reviewed by D Russell; comments to author 12.04.12; revised version received 18.04.12; accepted 26.04.12; published 13.06.12.

Please cite as:

Feufel MA, Stahl SF

What do Web-Use Skill Differences Imply for Online Health Information Searches?

J Med Internet Res 2012;14(3):e87

URL: <http://www.jmir.org/2012/3/e87/>

doi: [10.2196/jmir.2051](https://doi.org/10.2196/jmir.2051)

PMID: [22695686](https://pubmed.ncbi.nlm.nih.gov/22695686/)

©Markus A Feufel, S Frederica Stahl. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Parent Opinions About Use of Text Messaging for Immunization Reminders

Carolyn Rose Ahlers-Schmidt¹, PhD; Amy K Chesser², PhD; Angelia M Paschal³, PhD; Traci A Hart⁴, PhD; Katherine S Williams¹, MEd; Beryl Yaghmai¹, MD; Sapna Shah-Haque⁵, MD, MPH

¹University of Kansas School of Medicine - Wichita, Pediatrics, Wichita, KS, United States

²University of Kansas School of Medicine - Wichita, Family and Community Medicine, Wichita, KS, United States

³Mississippi University for Women, Health & Kinesiology, Columbus, MS, United States

⁴Wichita State University, Psychology, Wichita, KS, United States

⁵University of Kansas School of Medicine - Wichita, Internal Medicine, Wichita, KS, United States

Corresponding Author:

Carolyn Rose Ahlers-Schmidt, PhD

University of Kansas School of Medicine - Wichita

Pediatrics

1010 N. Kansas

Wichita, KS, 67214

United States

Phone: 1 316 293 1810

Fax: 1 316 293 2686

Email: cschmidt3@kumc.edu

Abstract

Background: Adherence to childhood immunization schedules is a function of various factors. Given the increased use of technology as a strategy to increase immunization coverage, it is important to investigate how parents perceive different forms of communication, including traditional means and text-message reminders.

Objective: To examine current forms of communication about immunization information, parents' satisfaction levels with these communication modes, perceived barriers and benefits to using text messaging, and the ideal content of text messages for immunization reminders.

Methods: Structured interviews were developed and approved by two Institutional Review Boards. A convenience sample of 50 parents was recruited from two local pediatric clinics. The study included a demographics questionnaire, the shortened form of the Test of Functional Health Literacy for Adults (S-TOFHLA), questions regarding benefits and barriers of text communication from immunization providers, and preferred content for immunization reminders. Content analyses were performed on responses to barriers, benefits, and preferred content (all Cohen's kappas > 0.70).

Results: Respondents were mostly female (45/50, 90%), white non-Hispanic (31/50, 62%), between 20–41 years (mean = 29, SD 5), with one or two children (range 1–9). Nearly all (48/50, 96%) had an S-TOFHLA score in the “adequate” range. All parents (50/50, 100%) engaged in face-to-face contact with their child's physician at appointments, 74% (37/50) had contact via telephone, and none of the parents (0/50, 0%) used email or text messages. Most parents were satisfied with the face-to-face (48/50, 96%) and telephone (28/50, 75%) communication. Forty-nine of the 50 participants (98%) were interested in receiving immunization reminders by text message, and all parents (50/50, 100%) were willing to receive general appointment reminders by text message. Parents made 200 comments regarding text-message reminders. Benefits accounted for 63.5% of comments (127/200). The remaining 37.5% (73/200) regarded barriers; however, no barriers could be identified by 26% of participants (13/50). Parents made 172 comments regarding preferred content of text-message immunization reminders. The most frequently discussed topics were date due (50/172, 29%), general reminder (26/172, 26%), and child's name (21/172, 12%).

Conclusions: Most parents were satisfied with traditional communication; however, few had experienced any alternative forms of communication regarding immunizations. Benefits of receiving text messages for immunization reminders far outweighed the barriers identified by parents. Few barriers identified were text specific. Those that were, centered on cost if parents did not have unlimited texting plans.

KEYWORDS

Provider-patient communication; child immunizations; text message

Introduction

Adherence to childhood immunization schedules is likely a function of various factors, including parents' health literacy skills [1-6], immunization knowledge [7,8], and perceived quality of patient-provider communication [9]. Other aspects of immunization communication, including parents' preferences and perceived barriers, may also influence adherence to childhood vaccination schedules.

Understanding immunization communication from the parents' perspective should help researchers and practitioners identify specific problems and needs, which might vary among groups. For instance, there is some indication of knowledge gaps among low-income mothers regarding vaccinations, the intended purposes of these immunizations, and appropriate vaccination schedules [7,8]. Determining what these and other parents' preferences are for receiving vaccination information, as well as identifying perceived communication barriers and needs, might help researchers and practitioners design more effective interventions.

It is possible that certain aspects of communication influence adherence to vaccination schedules. For instance, a recent study [10] investigated the relationship between parents' health literacy skills and their children's vaccination status. Early vaccination status did not show a significant association; however, the findings indicated a significant difference in vaccination status at 3 months and 7 months of age. These results highlight the importance of health literacy skills for subsequent vaccinations. Although beyond the scope of the study [10], questions remain about the extent that immunization communication (including appointment reminders) played a role, and whether specific aspects of immunization communication influenced the study results.

Traditionally, medical providers have used face-to-face interaction, pamphlets or handouts, and vaccination appointment reminders (eg, phone calls) as forms of immunization communication [11,12]. The use of technology to provide information, particularly vaccination appointment reminders, has increased over the past few years to include computer-generated auto-dialer phone calls, automated letters and postcards [11], and email reminder programs [12,13]. More recently with the passage of the Meaningful Use rule [14], medical providers with electronic health records (EHRs) are exploring issuing health maintenance reminders. Some success has been found with EHRs in capturing vaccination opportunities in a pediatric population [15] and issuing influenza reminders in older adults [16].

Another use of technology for communicating immunization information is mobile phone text-messaging systems, known as short message services (SMS). While mobile phones are a common commodity across age, gender, and socioeconomic groups [17], SMS-based interventions are in various stages of

development and use. Immunization reminders delivered by text-message interventions have shown promising results in some populations [18]. Although some studies have found support for text-message programs from parents of teens [19-22], others have found mixed reactions from medical practitioners [23] as well as organizational barriers and logistical issues that need to be addressed [24].

The impact of text-message reminders on adherence to childhood vaccination schedules continues to be explored. More information is needed to inform the researcher and practitioner of the parents' perspectives regarding this communication form, preferences they might have, and how text messaging might influence adherence to childhood vaccinations. The current study aims to address this gap.

The purpose of this study is to explore immunization communication utilizing text messages from the parents' perspective. Because it is critical that children are vaccinated as early as possible in order to avoid vaccine-preventable diseases [10], understanding what might work best from the parents' viewpoint would be helpful. The findings will contribute to timely and beneficial use of translational immunization and technology research to address adherence to childhood immunization schedules.

Methods

As part of a feasibility study to determine acceptance among parents of text messages for child immunization reminders, a series of parent interviews were developed and approved by two local Institutional Review Boards. The interview guide was developed by a team with expertise in the areas of pediatric infectious disease, health communication, health disparities, community psychology, and human factors psychology. The guide included informed consent, a 10-item demographic survey, a script of interview questions, and the shortened form of the Test of Functional Health Literacy for Adults (S-TOFHLA) [25].

A convenience sample of parents was recruited from two local group practice pediatric clinics. Parents were informed about the study by the receptionist or nurse at the clinic. If interested, they filled out a form with contact information. The research team conducted a preliminary phone screening of the parent to ensure eligibility criteria were met. Eligibility criteria included: (1) parent or caregiver of a child age < 2 years; (2) adult (≥ 18 years); (3) use a mobile phone for sending and receiving text messages; (4) English speaking; and (5) able to provide informed consent. Parents with valid contact information who met the inclusion criteria (71/95; 75%) were scheduled for a 30-45 minute in-person interview at the medical school. All interviews were scheduled within 1 week of the screening call. Of the 71 appointments made, 30% (21/71) were no-shows. All participants received a US \$35 gift card at the conclusion of the interview to cover time and travel costs.

After obtaining informed consent, a self-report demographic survey was administered. The 10 questions were in multiple-choice format, based on the Centers for Disease Control and Prevention's Behavioral Risk Factor Surveillance System (BRFSS) questionnaire [21]. Interview questions included experience level with technology and current usage patterns, willingness to receive text messages from their child's medical providers, and perceived barriers. All questions were open-ended, except relationship to technology (used a 6-point scale) and satisfaction with current communication types (used a 4-point scale). For these questions, participants were given a card with their response options.

Specific objectives were to examine the most prevalent forms of communication about immunization information, parents' satisfaction levels with these communication modes, and perceived barriers and benefits to using a novel form of communication—text messaging. Participants also provided content ideas for what the text messages should read. Following these interview activities, a follow-up study was performed to identify optimal text message content and comprehension of sample text messages. These procedures are described elsewhere [21].

Health Literacy Assessment

The S-TOFHLA [25] was administered to consenting parents/guardians. Participants completed the 36-item S-TOFHLA, which takes up to 7 minutes to administer and has both normal and large print versions. Scoring results range from 0 to 36; participants are categorized as having adequate health literacy if their S-TOFHLA score is 23–36, marginal health literacy if their score is 17–22, and inadequate health literacy if their score is 0–16. The S-TOFHLA consists of two sections: instructions for preparation for an upper gastrointestinal (GI) series and a Medicaid application [25]. The Gunning Fog readability levels are 4.3 and 10.4, respectively. The passages are set up using a modified Cloze procedure, where approximately every 6th word has been removed.

Statistical Analysis

All data were entered into Statistical Package for the Social Sciences (SPSS) 17.0. Frequencies and percentages are reported for categorical data; means and standard deviations are reported for continuous data. In addition, content analysis was performed on the open-ended item responses. Participant respondents were counted and quantified for reporting.

The two questions regarding barriers and benefits of text message-based reminders used emergent coding [26]. This method allows the content itself to determine the categories. Two investigators independently reviewed the content and compiled a list of emerging themes. Next, they compared lists and reconciled any differences in order to develop a final list. The investigators confirmed the reliability (Cohen's kappa > 0.70) of the results. For the question regarding advantages of using text messaging for reminders, 4 categories were identified and 4 categories were also identified for the question regarding barriers.

To assess responses to the question regarding the type of content text-message reminders should include, *a priori* categories

determined by an expert panel were used. There were 16 named content categories plus 1 “other” category. Participant comments were assigned to one of the 17 categories by two researchers independently coding the responses. The investigators confirmed the reliability (Cohen's kappa > 0.70). Discrepancies were corrected by consensus. Frequencies of individual categories were computed.

Results

Demographically, the majority of respondents were female (45/50, 90%), white non-Hispanic (31/50, 62%), married (20/50, 40%) or members of an unmarried couple (14/50, 28%), with one or two children (range 1–9). Participant age ranged from 20 to 41 years with a mean age of 29 years (SD 5). Most (28/50, 56%) completed 1–3 years of college and 40% (20/50) described their jobs as “employed for wages” with an annual income (from all sources) below US \$20,000 (30/50, 60%). Nearly all participants (48/50, 96%) had an S-TOFHLA score in the “adequate” range. One participant's health literacy level (1/50; 2%) was identified as “marginal” and another (1/50, 2%) had an “inadequate” score.

Regarding current communication with their child's physician, all (50/50, 100%) parents engaged in face-to-face contact at the appointments, 74% (37/50) reported communication via telephone, and none of the parents (0/50; 0%) reported using email or text communication. Parental satisfaction with each communication type is illustrated in Figure 1.

Parents reported obtaining the majority of information about immunizations for their child at doctor's appointments (39/50, 78%); in mailings from the Health Department, Women, Infants and Children program, or Medicaid (12/50, 24%); in mailings from their child's doctor's office (11/50, 22%); or the Internet (5/50, 10%). When asked how they know when it's time to schedule their child's immunizations, the majority of parents reported being told at their child's previous appointment and having to remember (38/50, 76%). For older children (> 1 year) some parents relied on memory cues based on their child's birthday or annual appointment (5/50, 10%) and others relied on the school to let them know (4/50, 8%). One father (1/50, 2%) admitted that his child's “mommy takes care of that” and if he were on his own he would have no idea what the schedule was.

Almost every participant (49/50, 98%) was interested in receiving immunization reminders by text message and 100% (50/50) were willing to receive general appointment reminders. In addition, 60% (30/50) would be willing to receive lab results by text—although several respondents wanted “only the good results by text, anything bad they should just, like send one that said ‘Lab results are in, please call.’ ” Other suggestions included alerts, such as “We now have the H1N1 vaccine available” (3/50, 6%) and follow-up texts after acute care (2/50, 4%).

For the content analysis, 200 individual comments were coded as relating to either benefits (127/200, 63.4%) or barriers (73/200, 36.5%). Parent comments suggesting benefits of text-message reminders fell into 4 emergent categories:

technology, convenience, communication, and general positive (see Table 1). The largest category of comments was “technology” (47/127, 37.0%). Many comments in this category addressed a dislike for talking on the phone or checking voicemail messages.

The barriers to text-message reminders had 4 emergent categories: technology, none identified, communication, and

other. “Technology” was also the largest category regarding barriers comments (43/73, 59%). Many of these comments addressed barriers such as if a phone was turned off or lost, while a few comments centered on cost if parents utilized pay-per-text programs. Over ¼ of participants (13/50, 26%) could not identify any barriers. The next most common barrier identified involved communication, with 9 of the 73 comments (12%).

Table 1. Content analysis of participant comments (N = 200) about benefits and barriers to text message reminders.

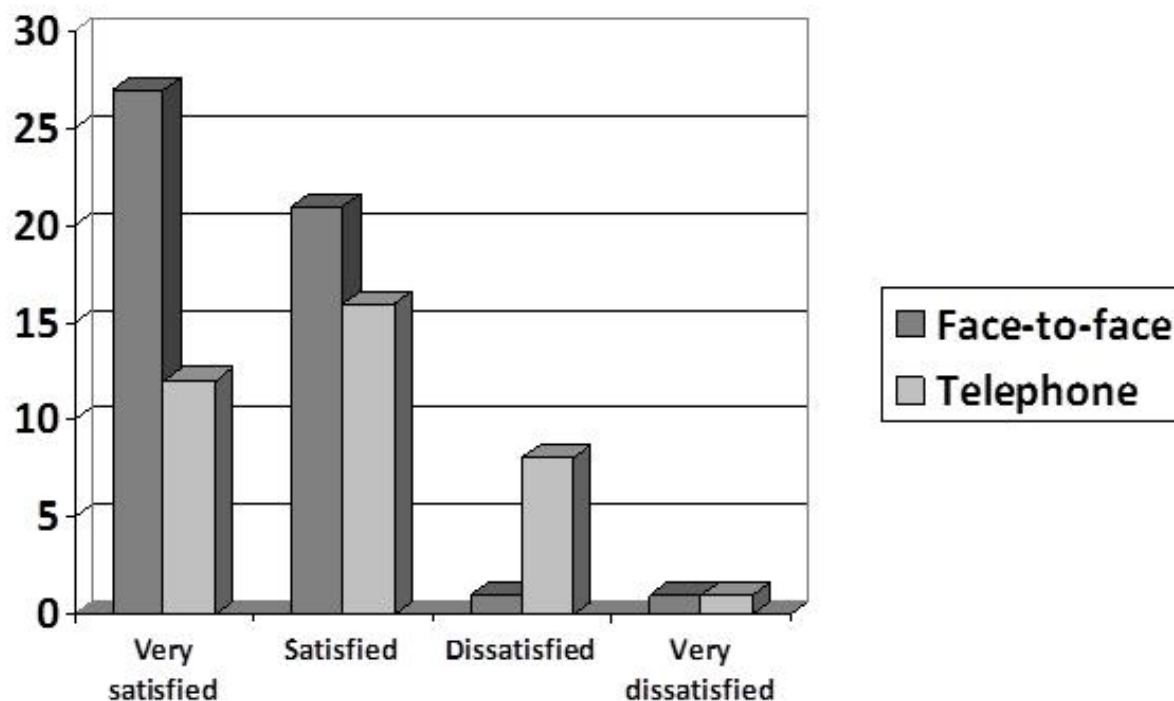
Category and emergent themes	Definition	Freq (%)	Examples of comments
Benefits			
Technology	The speed with which information is available, ability to link to other systems/calendars, or other comments related to cellular phone technology.	47 (37.0)	“Text doesn’t waste time and minutes like a phone call,” “You can put it right into appt book in phone,” and “Text will come through later if in bad reception where a phone call won’t.”
Convenience	Information related to the ease or timeliness of receiving reminders.	35 (27.6)	“Easier,” “quick,” “convenient,” and “I can lock it and go back and check it later.”
Communication	Includes preferences for, or avoidance of, specific avenues of communication.	23 (18.1)	“Saves time versus phone conversations,” “I respond more to a text,” and “I don’t always check missed calls, but I do check missed texts.”
General positive	Comments regarding the benefits or the usefulness of reminders in general.	22 (17.3)	“A reminder is good” and “I’m forgetful so I’ll have a written reminder.”
Barriers			
Technology	Comments related to costs, lack of text capabilities, or not being technology-savvy.	43 (58.9)	“You might get charged for text if you don’t have the unlimited plan,” “Turned off phones,” and “Text not working or delayed.”
None identified	Participant was unable to come up with any negatives to using text messaging.	13 (17.8)	“None,” “none for me,” or “no problems.”
Communication	Concerns regarding ability to understand text content, its limited characters, the use of abbreviations, or being provided inaccurate information.	9 (12.3)	“Accuracy,” “punctuation,” and “Parents might not understand the text if they don’t know a lot about immunizations.”
Other	Any item not appropriate for one of the above categories.	8 (11.0)	“Doctor’s office might get overwhelmed with texts coming in” and “forgetting about it.”

Finally, parents were asked to think about the content they would like to see in a text-message immunization reminder (see Table 2). These responses were categorized based on *a priori* categories. The most commonly identified information expected was “date immunizations are due.” The second most frequent response category was “other.” The majority of comments in

this category included suggestions of using simple, easy-to-understand language or suggestions of what “not” to include. The only item that fell into the other category that was informational content was doctor’s office location, which was identified by 10% (5/50) of parents.

Table 2. Frequency of participant comments (N = 172) about preferred text message immunization reminder information.

Category (<i>a priori</i>)	Example	Freq (%)
Date immunizations are due	“When appointment is,” “Appointment tomorrow at 9:15,” “When shots are due.”	50 (29.1)
Other information	“Location” and “Use numbers—9 months vs nine months.”	31 (18.0)
General reminder	“Immunization due soon” and “Time to schedule an appointment for shots.”	26 (15.1)
Child’s name	“Which child” and “Name of child.”	21 (12.2)
Name of vaccine	“What type of shot” and “Names of vaccines.”	16 (9.3)
Clinic number	“Phone number for the Doctor” and “If questions contact at this number.”	11 (6.4)
Child’s age	“Time frame—like 6 month immunizations” and “Baby’s age.”	9 (5.2)
Doctor’s name	“Doctor’s name” and “Dr. [name].”	3 (1.7)
Number of shots	“How many shots.”	2 (1.2)
Name of vaccine-preventable disease	“What the shot prevents.”	1 (0.6)
Side effects	“Side effects so you can be prepared for after the appointment.”	1 (0.6)
What the disease can do to my child	“Risks of the disease.”	1 (0.6)
Full immunization schedule		0 (0)
How the disease spreads		0 (0)
Total number of specific vaccine needed		0 (0)
Who is at risk?		0 (0)
Where your child is in the specific vaccine se- ries		0 (0)

Figure 1. Parent (N = 50) satisfaction with current communication methods with child's health care provider.

Discussion

The purpose of this study was to explore immunization communication among parents to better understand how adherence to childhood vaccinations can be addressed. Specifically, we wanted to identify the most prevalent forms of immunization communication, examine parents' satisfaction with these communication modes, and—because of the increased use of mobile phone text immunization reminders—explore perceived barriers and benefits to using text messaging.

Parents in our study were most likely to receive communication about immunization information face-to-face at the pediatricians' office. Most parents were satisfied with this form of communication; however, few had experienced any alternative forms of communication regarding immunizations. Parents indicated less satisfaction with telephone than face-to-face communication. Although several parents reported receiving mailers from their child's physician, no parents reported email or text message communication. This is consistent with results from a recent survey of physicians conducted in the same metropolitan area. In that study, Chesser et al [27] found that none of the responding pediatricians or family physicians were using email or text messaging to convey immunization reminders to parents; in addition, these physicians were divided in their opinions of the appropriateness of sending text messages to parents [23].

Parents, on the other hand, were interested in receiving text message communication from their child's physician. This may indicate that text messaging is considered a form of constructive communication from physicians, as suggested by Raine et al [28]. In fact, all participants were willing to receive text-message reminders for general appointments. One mother did not wish to receive immunization-specific text reminders; she reported her child was on an "alternative" immunization schedule, but would not share any more detail on what that meant.

Parents were able to identify nearly twice the benefits to receiving text messages as barriers, and most of the barriers identified were not text specific. For example, many comments addressed barriers such as if a phone was turned off or lost which would also apply to phone calls or voicemails to mobile phones. Most of the comments that could be applied exclusively to text messaging seemed centered on cost if parents did not have unlimited texting plans. However, these comments were meant generally and did not necessarily apply to the participant making them. Several parents who had unlimited text messaging capabilities stated, "It might get expensive for people who have to pay for each text." This barrier is not insurmountable as a recent survey of low-income parents found 81% of the 167 parents with text messaging had an unlimited text plan [19].

The second most common barrier identified was communication. This is not surprising as discrepancies between physicians and patients in general literacy levels can result in misunderstanding of written handouts and have resulted in efforts to rewrite materials to improve their usefulness to patients [29]. Whereas health literacy is one explanation for poor patient understanding of medical information, determining who has low health literacy presents a challenge to physicians. Physicians can enhance

understanding by assessing health literacy at new patient visits, refraining from use of medical terminology, drawing a picture, limiting information, assessing understanding using a teach-back method, and being respectful and sensitive in order to empower patients [30,31].

In the current study, we first asked parents to free-associate the content they would like to see in a text-message reminder and later, in the follow-up study, we asked them to organize cards with sample content into critical, clarifying, or unnecessary information to include in a text message [21]. The top 6 critical content items selected from the card sort were in the same categories mentioned by interviewed parents as content they would like to see in a text-message immunization reminder. This suggests some reliability in the responses given by parents. In addition, these results were consistent with other published studies [20,22]. As with these previous studies, participants in the current study suggested that the information be short, simple, and personalized. One item we had not anticipated was the importance of the clinic address to parents. This information may be most important to include in reminders for the initial immunization appointment or for certain populations and should be further investigated.

It is important to remember that the limitation of text messages to 160 characters also limits the information that can be conveyed with this technology. While text-message immunization reminders may have the capability to enhance parental understanding of immunization schedules, this is probably not the most appropriate mode to address parents' attitudes, knowledge, and trust regarding vaccines. Wu et al identified knowledge gaps in mothers' understanding of vaccinations and suggested mothers would benefit from discussions regarding risks and benefits of vaccines during prenatal care [32].

Limitations

While this study was important for understanding immunization communication from parents' viewpoints, there are limitations that should be considered. First, self-reported preferences, behaviors, and experiences might be an issue as some parents may have inaccurately responded to interview questions. Socially desirable responses might have impacted the results.

Generalizability of the study findings might also be an issue due to the small sample size. Participating parents came from a single Midwestern urban city, had relatively high education and health literacy levels, spoke only English, and were proficient with text messaging. Thus, participants in this study might not have been representative of most parents, even in the targeted city. Further research is needed to examine immunization communication in larger samples and among diverse groups of parents. Results might differ among those with less text messaging experience, marginal or low health literacy, and varied ethnicity [20] and language.

Conclusions

The goal of this study was to provide information that will lead to the development of parental interventions and programs to help improve adherence to childhood immunization schedules. The research findings are being used to design an intervention

versus control model to be used in a future study assessing impact of text messages on vaccine schedule compliance. Based on the premise that children's health is influenced by various parental factors, including parents' health literacy skills, knowledge, and perceptions of patient-provider communication about immunization information, these results may be used to help providers and researchers effectively promote early adherence, services, and immunization programs directly to parents.

Practice Implications

While other studies have assessed whether various types of immunization reminders improve immunization rates [33] more information is needed. Limited information existed regarding what communication methods were preferred and what issues might exist according to parents. With the increased use of text messaging as a form of communication, it was important to

know whether this was a viable form of immunization communication with parents. The current study addressed these needs and provided greater understanding of immunization communication from the parents' perspective. The findings provide a basis for further research and text-message interventions that could address adherence to childhood immunization schedules.

The findings of this study and others [19-22] suggest an overwhelming support for text-message immunization reminders from parents who utilize text messaging. Text messaging may not only be a viable tool for increasing immunization compliance in children, but may actually be the preferred tool of some subgroups of parents. In general, physicians should consider utilizing advances in SMS-based technology to enhance communication regarding immunizations—a benefit seen in other health promotion programs including diabetes management [34] and weight loss [35].

Acknowledgments

We would like to thank Dr. Mark Russell, Dr. William Hou, and Karen Davis, MA, for their hard work on this project. This project was funded by a University of Kansas School of Medicine-Wichita Level I Dean's Fund Grant Award.

Conflicts of Interest

None declared.

References

1. Alderson P, Sutcliffe K, Curtis K. Children as partners with adults in their medical care. *Arch Dis Child* 2006 Apr;91(4):300-303. [doi: [10.1136/adc.2005.079442](https://doi.org/10.1136/adc.2005.079442)] [Medline: [16399782](https://pubmed.ncbi.nlm.nih.gov/16399782/)]
2. Butz AM, Walker JM, Pulsifer M, Winkelstein M. Shared decision making in school age children with asthma. *Pediatr Nurs* 2007;33(2):111-116. [Medline: [17542232](https://pubmed.ncbi.nlm.nih.gov/17542232/)]
3. Lokker N, Sanders L, Perrin EM, Kumar D, Finkle J, Franco V, et al. Parental misinterpretations of over-the-counter pediatric cough and cold medication labels. *Pediatrics* 2009 Jun;123(6):1464-1471 [FREE Full text] [doi: [10.1542/peds.2008-0854](https://doi.org/10.1542/peds.2008-0854)] [Medline: [19482755](https://pubmed.ncbi.nlm.nih.gov/19482755/)]
4. Sanders LM, Federico S, Klass P, Abrams MA, Dreyer B. Literacy and child health: a systematic review. *Arch Pediatr Adolesc Med* 2009 Feb;163(2):131-140 [FREE Full text] [doi: [10.1001/archpediatrics.2008.539](https://doi.org/10.1001/archpediatrics.2008.539)] [Medline: [19188645](https://pubmed.ncbi.nlm.nih.gov/19188645/)]
5. Yin HS, Johnson M, Mendelsohn AL, Abrams MA, Sanders LM, Dreyer BP. The health literacy of parents in the United States: a nationally representative study. *Pediatrics* 2009 Nov;124 Suppl 3:S289-S298 [FREE Full text] [doi: [10.1542/peds.2009-1162E](https://doi.org/10.1542/peds.2009-1162E)] [Medline: [19861483](https://pubmed.ncbi.nlm.nih.gov/19861483/)]
6. Yin HS, Mendelsohn AL, Wolf MS, Parker RM, Fierman A, van Schaick L, et al. Parents' medication administration errors: role of dosing instruments and health literacy. *Arch Pediatr Adolesc Med* 2010 Feb;164(2):181-186 [FREE Full text] [doi: [10.1001/archpediatrics.2009.269](https://doi.org/10.1001/archpediatrics.2009.269)] [Medline: [20124148](https://pubmed.ncbi.nlm.nih.gov/20124148/)]
7. Evers DB. Teaching mothers about childhood immunizations. *MCN Am J Matern Child Nurs* 2001;26(5):253-256. [Medline: [11552575](https://pubmed.ncbi.nlm.nih.gov/11552575/)]
8. Baker LM, Wilson FL, Nordstrom CK, Legwand C. Mothers' knowledge and information needs relating to childhood immunizations. *Issues Compr Pediatr Nurs* 2007;30(1-2):39-53. [doi: [10.1080/01460860701366666](https://doi.org/10.1080/01460860701366666)] [Medline: [17613141](https://pubmed.ncbi.nlm.nih.gov/17613141/)]
9. Hironaka LK, Paasche-Orlow MK. The implications of health literacy on patient-provider communication. *Arch Dis Child* 2008 May;93(5):428-432. [doi: [10.1136/adc.2007.131516](https://doi.org/10.1136/adc.2007.131516)] [Medline: [17916588](https://pubmed.ncbi.nlm.nih.gov/17916588/)]
10. Pati S, Feemster KA, Mohamad Z, Fiks A, Grundmeier R, Cnaan A. Maternal health literacy and late initiation of immunizations among an inner-city birth cohort. *Matern Child Health J* 2011 Apr;15(3):386-394. [doi: [10.1007/s10995-010-0580-0](https://doi.org/10.1007/s10995-010-0580-0)] [Medline: [20180003](https://pubmed.ncbi.nlm.nih.gov/20180003/)]
11. Franzini L, Rosenthal J, Spears W, Martin HS, Balderas L, Brown M, et al. Cost-effectiveness of childhood immunization reminder/recall systems in urban private practices. *Pediatrics* 2000 Jul;106(1 Pt 2):177-183. [Medline: [10888689](https://pubmed.ncbi.nlm.nih.gov/10888689/)]
12. Luthy KE, Sperhac AM, Faux SA, Miner JK. Modern Medicine. 2010. Improving immunization rates in the clinic and in the community URL: <http://www.modernmedicine.com/modernmedicine/Modern+Medicine+Now/Improving-immunization-rates-in-the-clinic-and-in-/ArticleStandard/Article/detail/687345> [accessed 2011-10-17] [WebCite Cache ID 62VZDMwTK]

13. Utah Department of Health, Immunization Program. 2002. Electronic Immunization Reminder Service URL: <http://www.immunize-utah.org/apps/uir/index.php> [accessed 2011-10-17] [WebCite Cache ID 62VcdK2Rn]
14. H.R. 1--111th Congress: American Recovery and Reinvestment Act of 2009. 2009 Jan 26. URL: <http://www.govtrack.us/congress/bills/111/hr1> [accessed 2012-05-10] [WebCite Cache ID 67YgICQRJ]
15. Fiks AG, Grundmeier RW, Biggs LM, Localio AR, Alessandrini EA. Impact of clinical alerts within an electronic health record on routine childhood immunization in an urban pediatric population. *Pediatrics* 2007 Oct;120(4):707-714 [FREE Full text] [doi: [10.1542/peds.2007-0257](https://doi.org/10.1542/peds.2007-0257)] [Medline: [17908756](https://pubmed.ncbi.nlm.nih.gov/17908756/)]
16. Ledwich LJ, Harrington TM, Ayoub WT, Sartorius JA, Newman ED. Improved influenza and pneumococcal vaccination in rheumatology patients taking immunosuppressants using an electronic health record best practice alert. *Arthritis Rheum* 2009 Nov 15;61(11):1505-1510 [FREE Full text] [doi: [10.1002/art.24873](https://doi.org/10.1002/art.24873)] [Medline: [19877088](https://pubmed.ncbi.nlm.nih.gov/19877088/)]
17. Blumberg SJ, Luke JV, Davidson G, Davern ME, Yu TC, Soderberg K. Wireless substitution: state-level estimates from the National Health Interview Survey, January-December 2007. *Natl Health Stat Report* 2009 Mar 11(14):1-13, 16. [Medline: [19418704](https://pubmed.ncbi.nlm.nih.gov/19418704/)]
18. Kharbanda EO, Stockwell MS, Fox HW, Andres R, Lara M, Rickert VI. Text message reminders to promote human papillomavirus vaccination. *Vaccine* 2011 Mar 21;29(14):2537-2541. [doi: [10.1016/j.vaccine.2011.01.065](https://doi.org/10.1016/j.vaccine.2011.01.065)] [Medline: [21300094](https://pubmed.ncbi.nlm.nih.gov/21300094/)]
19. Ahlers-Schmidt CR, Chesser A, Hart T, Paschal A, Nguyen T, Wittler RR. Text messaging immunization reminders: feasibility of implementation with low-income parents. *Prev Med* 2010;50(5-6):306-307. [doi: [10.1016/j.ypmed.2010.02.008](https://doi.org/10.1016/j.ypmed.2010.02.008)] [Medline: [20178813](https://pubmed.ncbi.nlm.nih.gov/20178813/)]
20. Ahlers-Schmidt CR, Hart T, Chesser A, Paschal A, Nguyen T, Wittler RR. Content of text messaging immunization reminders: what low-income parents want to know. *Patient Educ Couns* 2011 Oct;85(1):119-121. [doi: [10.1016/j.pec.2010.08.007](https://doi.org/10.1016/j.pec.2010.08.007)] [Medline: [20832970](https://pubmed.ncbi.nlm.nih.gov/20832970/)]
21. Ahlers-Schmidt CR, Hart T, Chesser A, Williams KS, Yaghamai B, Shah-Haque S, et al. Using Human Factors Techniques to Design Text Message Reminders for Childhood Immunization. *Health Educ Behav* 2011 Oct 10. [doi: [10.1177/1090198111420866](https://doi.org/10.1177/1090198111420866)] [Medline: [21986245](https://pubmed.ncbi.nlm.nih.gov/21986245/)]
22. Kharbanda EO, Stockwell MS, Fox HW, Rickert VI. Text4Health: a qualitative evaluation of parental readiness for text message immunization reminders. *Am J Public Health* 2009 Dec;99(12):2176-2178. [doi: [10.2105/AJPH.2009.161364](https://doi.org/10.2105/AJPH.2009.161364)] [Medline: [19833982](https://pubmed.ncbi.nlm.nih.gov/19833982/)]
23. Hart T, Ahlers-Schmidt CR, Chesser A, Jones J, Williams KS, Wittler RR. Physician impressions of using text message technology to increase vaccination compliance. *Telemed J E Health* 2011;17(6):427-430. [doi: [10.1089/tmj.2010.0221](https://doi.org/10.1089/tmj.2010.0221)] [Medline: [21612518](https://pubmed.ncbi.nlm.nih.gov/21612518/)]
24. Karasz H, Bogan S. What 2 know b4 u text: Short Message Service options for local health departments. *Washington State Journal of Public Health Practice* 2011;4(1):20-27 [FREE Full text]
25. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. *Patient Educ Couns* 1999 Sep;38(1):33-42. [Medline: [14528569](https://pubmed.ncbi.nlm.nih.gov/14528569/)]
26. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
27. Chesser A, Paschal A, Hart T, Jones J, Williams KS, Ahlers-Schmidt CR. Communication practices for pediatric immunization information: physician perceptions of parent health literacy skills. *Clin Pediatr (Phila)* 2012 May;51(5):504-506. [doi: [10.1177/0009922810396548](https://doi.org/10.1177/0009922810396548)] [Medline: [21525085](https://pubmed.ncbi.nlm.nih.gov/21525085/)]
28. Raine R, Cartwright M, Richens Y, Mahamed Z, Smith D. A qualitative study of women's experiences of communication in antenatal care: identifying areas for action. *Matern Child Health J* 2010 Jul;14(4):590-599. [doi: [10.1007/s10995-009-0489-7](https://doi.org/10.1007/s10995-009-0489-7)] [Medline: [19554436](https://pubmed.ncbi.nlm.nih.gov/19554436/)]
29. Plimpton S, Root J. Materials and strategies that work in low literacy health communication. *Public Health Rep* 1994;109(1):86-92. [Medline: [8303020](https://pubmed.ncbi.nlm.nih.gov/8303020/)]
30. Simpson M, Buckman R, Stewart M, Maguire P, Lipkin M, Novack D, et al. Doctor-patient communication: the Toronto consensus statement. *BMJ* 1991 Nov 30;303(6814):1385-1387. [Medline: [1760608](https://pubmed.ncbi.nlm.nih.gov/1760608/)]
31. Williams MV, Davis T, Parker RM, Weiss BD. The role of health literacy in patient-physician communication. *Fam Med* 2002 May;34(5):383-389. [Medline: [12038721](https://pubmed.ncbi.nlm.nih.gov/12038721/)]
32. Wu AC, Wisler-Sher DJ, Griswold K, Colson E, Shapiro ED, Holmboe ES, et al. Postpartum mothers' attitudes, knowledge, and trust regarding vaccination. *Matern Child Health J* 2008 Nov;12(6):766-773. [doi: [10.1007/s10995-007-0302-4](https://doi.org/10.1007/s10995-007-0302-4)] [Medline: [17987370](https://pubmed.ncbi.nlm.nih.gov/17987370/)]
33. Jacobson VJ, Szilagyi P. Patient reminder and patient recall systems to improve immunization rates. *Cochrane Database Syst Rev* 2005(3):CD003941. [doi: [10.1002/14651858.CD003941.pub2](https://doi.org/10.1002/14651858.CD003941.pub2)] [Medline: [16034918](https://pubmed.ncbi.nlm.nih.gov/16034918/)]
34. Franklin VL, Greene A, Waller A, Greene SA, Pagliari C. Patients' engagement with "Sweet Talk" - a text messaging support system for young people with diabetes. *J Med Internet Res* 2008;10(2):e20 [FREE Full text] [doi: [10.2196/jmir.962](https://doi.org/10.2196/jmir.962)] [Medline: [18653444](https://pubmed.ncbi.nlm.nih.gov/18653444/)]

35. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. *J Med Internet Res* 2009;11(1):e1 [FREE Full text] [doi: [10.2196/jmir.1100](https://doi.org/10.2196/jmir.1100)] [Medline: [19141433](https://pubmed.ncbi.nlm.nih.gov/19141433/)]

Abbreviations

BRFSS: Behavioral Risk Factor Surveillance System

GI: gastrointestinal

SMS: short message service

SPSS: Statistical Package for the Social Sciences

S-TOFHLA: Test of Functional Health Literacy for Adults (shortened form)

Edited by G Eysenbach; submitted 28.10.11; peer-reviewed by M Stockwell, R Furberg; comments to author 11.01.12; revised version received 10.02.12; accepted 09.03.12; published 06.06.12.

Please cite as:

Ahlers-Schmidt CR, Chesser AK, Paschal AM, Hart TA, Williams KS, Yaghmai B, Shah-Haque S

Parent Opinions About Use of Text Messaging for Immunization Reminders

J Med Internet Res 2012;14(3):e83

URL: <http://www.jmir.org/2012/3/e83/>

doi: [10.2196/jmir.1976](https://doi.org/10.2196/jmir.1976)

PMID: [22683920](https://pubmed.ncbi.nlm.nih.gov/22683920/)

©Carolyn Rose Ahlers-Schmidt, Amy K Chesser, Angelia M Paschal, Traci A Hart, Katherine S Williams, Beryl Yaghmai, Sapna Shah-Haque. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 06.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Disseminating Self-Help: Positive Psychology Exercises in an Online Trial

Stephen M Schueller^{1,2}, PhD; Acacia C Parks^{1,3}, PhD

¹University of Pennsylvania, Department of Psychology, Positive Psychology Center, Philadelphia, PA, United States

²University of California, San Francisco/San Francisco General Hospital, Department of Psychiatry, San Francisco, CA, United States

³Hiram College, Department of Psychology, Hiram, OH, United States

Corresponding Author:

Stephen M Schueller, PhD

University of California, San Francisco/San Francisco General Hospital

Department of Psychiatry

1001 Potrero Avenue, Suite 7M

San Francisco, CA, 94110

United States

Phone: 1 415 206 5279

Fax: 1 415 206 8942

Email: stephen.schueller@ucsf.edu

Abstract

Background: The recent growth of positive psychology has led to a proliferation in exercises to increase positive thoughts, behaviors, and emotions. Preliminary evidence suggests that these exercises hold promise as an approach for reducing depressive symptoms. These exercises are typically researched in isolation as single exercises. The current study examined the acceptability of several multi-exercise packages using online dissemination.

Objective: The purpose of this study was to investigate methods of dissemination that could increase the acceptability and effectiveness of positive psychology exercises. To achieve this goal, we compared the use of positive psychology exercises when delivered in packages of 2, 4, or 6 exercises.

Methods: Self-help-seeking participants enrolled in this study by visiting an online research portal. Consenting participants were randomly assigned to receive 2, 4, or 6 positive psychology exercises (or assessments only) over a 6-week period. These exercises drew from the content of group positive psychotherapy. Participants visited an automated website that distributed exercise instructions, provided email reminders, and contained the baseline and follow-up assessments. Following each exercise, participants rated their enjoyment of the exercise, answered how often they had used each technique, and completed outcome measures.

Results: In total, 1364 individuals consented to participate. Attrition rates across the 2-, 4-, and 6-exercise conditions were similar at 55.5% (181/326), 55.8% (203/364), and 52.7% (168/319) respectively but were significantly greater than the attrition rate of 42.5% (151/355) for the control condition ($\chi^2_3 = 16.40$, $P < .001$). Participants in the 6-exercise condition were significantly more likely than participants in the 4-exercise condition to use both the third ($F_{1,312} = 5.61$, $P = .02$) and fourth ($F_{1,313} = 6.03$, $P = .02$) exercises. For 5 of the 6 exercises, enjoyment was related to continued use of the exercise at 6-week follow-up (r 's = .12 to .39). All conditions produced significant reductions in depressive symptoms ($F_{1,656} = 94.71$, $P < .001$); however, a significant condition by time interaction ($F_{3,656} = 4.77$, $P = .003$) indicated that this reduction was larger in the groups that received 2 or 4 exercises compared with the 6-exercise or control condition.

Conclusion: Increasing the number of exercises presented to participants increased the use of the techniques and did not increase dropout. Participants may be more likely to use these skills when presented with a variety of options. Increasing the number of exercises delivered to participants produced a curvilinear relationship with those in the 2- and 4-exercise conditions reporting larger decreases in depressive symptoms than participants in the 6-exercise or control conditions. Although research generally offers a single exercise to test isolate effects, this study supports that studying variability in dissemination can produce important findings.

(*J Med Internet Res* 2012;14(3):e63) doi:[10.2196/jmir.1850](https://doi.org/10.2196/jmir.1850)

KEYWORDS

Internet intervention; positive psychology; self-help

Introduction

Depression is a major global public health issue affecting approximately 121 million people worldwide, a higher prevalence than any other mental disorder [1]. Costing an estimated US \$43 billion per year in treatment, depression is the leading cause of disability and lost productivity worldwide [2]. Furthermore, depression is a chronic disease; once an individual has experienced his or her first episode, he or she is 2 to 3 times more likely to experience subsequent episodes [3]. The extent of the depression epidemic is furthered by an additional proportion of the population reporting depressive symptoms that do not meet criteria for a major depressive episode. Mild to moderate depressive symptoms are at least as prevalent as major depressive disorder, and many individuals with subthreshold depressive symptoms experience substantial functional impairment [4,5]. Furthermore, subthreshold symptoms are a strong risk factor for developing major depression in the future [6].

A growing awareness of the burden caused by major depression and impairment due to subclinical symptoms has fostered interest in strategies that promote mental health in individuals before they have experienced full episodes of major depression. One promotion strategy, highlighted by positive psychology, is to improve mental health by increasing the frequencies of positive emotions, behaviors, and cognitions rather than targeting negative emotions and thoughts. Furthermore, the burden of depression might also be reduced through developing nontraditional methods of disseminating mental health treatments to increase access to these treatments. The Internet is one such method to help deliver empirically validated treatments to a broader audience. This study combines these strategies and nontraditional methods of dissemination into a single investigation of positive psychology techniques disseminated through an Internet-based system.

The Potential of Positive Psychology

Positive psychology focuses on factors that help promote optimal well-being and investigates the functional consequences of these states [7]. A recent theory by Seligman proposes five pathways to well-being: positive emotions, engagement, relationships, meaning, and achievement [8]. Seligman posits that these pathways are unique predictors of well-being, each of which people pursue and value in their own right.

However, these pathways are linked not only to higher levels of self-reported psychological well-being but also to tangible functional benefits. Experimental and longitudinal studies have found that increased positive emotions lead to social, occupational, and health benefits [9]. Gratitude, for example, promotes both the formation and the upkeep of friendships [10]. Correlational evidence links higher reported focus on engaging and meaningful activities to greater occupational success and higher incomes [11]. Consistent findings suggest that stronger social relationships reduce mortality and help us cope with life stressors [12].

Furthermore, psychological well-being in general, and positive emotion in particular, appear to promote meaning-making in stressful situations, which in turn leads to more effective coping [13]. In a study examining the role of positive emotions in psychological resilience, individuals who experienced positive emotions in the weeks immediately following September 11th were less likely to experience increased depressive symptom levels, controlling for baseline symptoms [14]. Even more interestingly, a recent study showed that increased psychological well-being is a significant protective factor from the occurrence of future episodes of major depression [15]. Indeed, this effect remained even when controlling for baseline levels of depressive symptoms, suggesting that positive mental health is different from merely the absence of psychological distress [16]. Thus, increasing well-being might not only increase positive factors but also reduce negative states and psychopathology.

Positive Psychology Interventions

A variety of researchers and practitioners have translated theories and findings based on positive emotions, well-being, and interpersonal functioning (ie, gratitude, kindness, positive responding, and savoring) into cognitive and behavioral practices deemed positive interventions. The most common outcome measures used to evaluate these interventions reflect aspects of subjective well-being and include increasing cognitive evaluations (such as life satisfaction or happiness), promoting positive emotions, and decreasing negative emotions (or reducing depressive symptoms).

In one of the first empirical investigations of an intervention aimed to increase happiness, Fordyce provided college students with a course that taught participants to mimic the behaviors of happy people [17]. These skills included behavioral recommendations such as increasing physical and social activity and cognitive strategies such as reducing one's expectations and cultivating a present-focused orientation. This course led to significant improvements in happiness compared with a placebo control condition. More recent research has focused on isolating individual factors and delivering single interventions. A selection of individual exercises with various targets (gratitude, strengths, and reminiscence) are reviewed below.

Gratitude

Gratitude is an emotion related to the reflection on something good that has happened and an acknowledgment of who or what was responsible. As previously mentioned, higher levels of gratitude are linked to improved relationships and coping with stressors [10,14]. One exercise designed to increase gratitude asks participants to reflect each day on three good things that happened by writing these things down in a journal [18]. College students who noted good things reported increased levels of positive affect, increased levels of exercise, better sleep quality, and fewer physical symptoms compared with a control condition where participants wrote about daily hassles. A further investigation of gratitude journaling found that participants benefited more when completing longer weekly entries as opposed to shorter more frequent entries [19]. These results

suggest that although even brief techniques can benefit individuals, the strategies must push participants to make large enough changes to be different than their typical routine.

One study comparing five different positive psychology exercises head-to-head used two different gratitude-boosting strategies [20]. One was a replication of gratitude journaling that also asked participants to reflect on the cause of the positive event, and the other required participants to write a gratitude letter to someone that they never had a proper chance to thank and to schedule a “gratitude visit” to read the letter to the respondent. Participants completed each exercise for one week and were followed for six months afterwards. The gratitude visit produced the largest initial gains in happiness; however, at three months, these participants were no happier than they had been at baseline. The gratitude journaling or the three good things exercise did not display benefits for participants until one month after the intervention, but these participants maintained these gains at 3- and 6-month follow-ups.

Strengths

In addition to examining the benefits of gratitude, the previous study also examined the benefits of encouraging participants to use one of their “signature strengths.” A signature strength is based on a classification of positive characteristics of individuals that are morally valued in general, that are fulfilling, and that are trait-like [21]. In this exercise, participants complete an evaluation of their strengths that provides them with feedback as to the ranking of their individual strength profile that highlights their top five signature strengths. Participants who practiced using their signature strengths each day showed increase happiness and decreased depressive symptoms compared with participants in a placebo control group that wrote about early memories.

Reminiscence

Similar to gratitude journaling, which encourages participants to identify and note good things during each day, looking back on one’s life and reflecting on positive moments is another strategy to promote well-being. Reminiscence interventions lead to a variety of emotional benefits, especially among elderly individuals. People who tended to reflect on the positive moments of their life reported increased ability to savor life and had higher levels of positive emotions [22]. A program that encouraged people to set aside 10 minutes twice a day for one week to reflect on positive moments in their life showed that this led to improved mood compared with individuals who thought about current interests or concerns [22].

The Efficacy of Positive Interventions

Thus, various strategies can help promote individual well-being. A recent meta-analysis demonstrated that compared with control conditions, positive interventions lead to reliable and moderate boosts in well-being and improvements in depressive symptoms [23]. Although many of the interventions were brief, single intervention techniques, moderation analysis revealed higher effect sizes associated with “shotgun” approaches, that is, packages that incorporate multiple positive psychology strategies. In light of this, the way to produce the most efficacious intervention might be to combine techniques from

previously validated individual exercises into an intervention package.

Positive psychotherapy (PPT) adopted this approach by drawing content from single exercises to create individual and group formats delivered by a single clinician [24]. PPT is a novel and distinct therapy technique that focuses on positive emotions, engagement, and meaning rather than targeting negative states and depressive symptoms. Individual PPT was piloted in a 14-session model for patients with a major depressive disorder recruited from a university counseling center. In a 12-week trial, individual PPT compared with eclectic counseling or eclectic counseling plus medication led to greater improvements in symptoms of depression at the end of treatment [24]. Group PPT was tested in a 6-week, 6-session model for groups of 8 to 11 mild-to-moderately depressed university students. Compared with students assigned to an assessment-only control group, students assigned to group PPT experienced significantly greater symptom reduction and improvements in life satisfaction, and these improvements remained at the 1-year follow-up [24]. These studies indicate the feasibility of these techniques to reduce depressive symptoms in individuals with clinical and subclinical levels of depression. The design of the investigation that we report in this paper translated the individual exercises used in group PPT for online dissemination.

Other programs that have combined content from individual exercises into longer and more involved intervention packages have been delivered in a variety of settings including schools and have also been delivered to military personnel [25,26]. Although outcome studies of these programs are still pending, the reaction from participants has been extremely positive, illustrating the widespread appeal of these techniques. Disseminating these programs, however, is limited by the need to train personnel and the ability of the programs to reach others. A meta-analysis of the school-based Penn Resiliency Program found that only trials delivered by highly trained research team providers produced significant gains [27]. Improving the dissemination of these techniques should be a focus of future research. The study we report here was intended to contribute to this goal.

Taking Positive Psychology Online

The brief duration and universal appeal of these strategies make positive psychology interventions ideal candidates for online administration. In fact, one of the first empirical tests of positive psychology interventions was conducted via the Internet using a convenience sample drawn from visitors to the website linked to Seligman’s book *Authentic Happiness* [28]. The Internet allows dissemination of these techniques throughout the world, anywhere participants have Web access, rather than having to rely on trained professionals to provide access to these techniques. Positive psychology techniques also lend themselves to Internet research, as these techniques might be less stigmatizing than other interventions that deal with difficulties and deficits. In positive psychology exercises, participants talk about the best aspects of themselves, reflect on what is going well in their lives, and reach out to others in their social network in appreciation. The three good things exercise was even integrated into Facebook through an application that provided

users with tools to complete the gratitude journaling online and share with friends in their network [29]. Although the application was not evaluated for efficacy, availability through a Facebook application did lead to increased usage. People using the application on Facebook were roughly twice as likely to report good things compared with users of a three good things Facebook *group*, which is distinct from an application in that it is merely a place where people can post and has no features [29].

Researchers have also used other platforms to promote dissemination of positive psychology interventions. Several individual exercises have been incorporated into *Live Happy*, a positive psychology app developed for the iPhone. In a naturalistic evaluation, participants who used the app reported increases in their mood and happiness throughout a 2-week period [30]. Participants who completed more activities within the app reported larger boosts in well-being than those who used the app less frequently.

Positive psychology strategies have also been used to augment other interventions to increase usage. In a study of an online body dissatisfaction intervention, participants who completed a gratitude diary were twice as likely to complete the intervention compared with a thought-monitoring and cognitive-restructuring group [31]. This relates to the theory underlying PPT, which contends that focusing on strengths and positive emotions fosters engagement and meaning more so than other techniques. Positive psychology techniques might be more enjoyable and acceptable than other interventions. This

is consistent with the overwhelming positive feedback from military and school-based positive psychology programs.

Current Study

Thus far, we have discussed existing positive psychology approaches and two important limitations of these approaches: feasibility (most interventions are administered in person, often by a therapist) and length (most interventions consist of a single exercise). The goal of the current study was to address both of these issues by offering a Web-based intervention that would last between 2 and 6 weeks. We aimed to systematically examine the effects of increased intervention length on adherence and efficacy, while also gaining basic usability data about the acceptability of an online positive psychology intervention that offers more than a single exercise.

The current study adapted group PPT into an online environment in order to investigate these aims. As previously stated, a meta-analysis of existing positive interventions demonstrated that longer interventions and packages with several techniques typically reported larger effect sizes. Our Web-based trial of positive interventions was designed to investigate these questions by comparing sets of 2, 4, or 6 exercises as well as an assessment-only control condition in terms of usage of the techniques (attrition and self-report) and efficacy (decreases in depressive symptoms). Exercises were drawn from group PPT and included active-constructive responding, a gratitude visit, life summary, three good things, savoring, and strengths. Table 1 provides brief descriptions of each of the exercises.

Table 1. Descriptions of individual exercises

Name of Exercise	Description of Exercise	Empirical Support (authors, publication year)
Active-constructive responding	Participants learn to respond positively to good news shared by others by lengthening the conversation and helping the sharer relive the experience.	Gable et al, 2004 [32]
Gratitude visit	Participants write a letter of gratitude and read it aloud to the target of the letter.	Seligman et al, 2005 [20]; Seligman et al, 2006 [24]
Life summary	Participants write a summary of how they would want their life expressed to their progeny.	Seligman et al, 2005 [20]; Seligman et al, 2006 [24]
Three good things	Participants identify three things that went well each day and why. These good things are kept in a gratitude journal throughout the week.	Emmons and McCullough, 2003 [18]; Seligman et al, 2005 [20]
Savoring	Participants are instructed to take time to focus intently on a positive experience 2 to 3 times each day.	Seligman et al, 2006 [24]
Strengths	Participants take the Values in Action Strengths Questionnaire and receive individualized feedback about their strengths and are instructed to use one of their top five (signature) strengths each day.	Seligman et al, 2005 [20]; Seligman et al, 2006 [24]

We posited that similar to the in-person group PPT, the online administration would lead to significantly greater decreases in depressive symptoms. Furthermore, consistent with the findings of the meta-analysis of positive interventions, we predicted that participants receiving more exercises would experience greater benefits.

Methods

Recruitment

Participants were recruited through a research portal associated with the University of Pennsylvania's Positive Psychology Center and Seligman's book *Authentic Happiness* [28]. A link available on the websites www.authentichappiness.org and www.ppresearch.sas.upenn.edu invited interested individuals to participate in a research study on positive psychology exercises.

Study Procedures

These landing sites offered participants a variety of assessment and research opportunities. From the initial website, individuals could access and receive feedback on questionnaires assessing emotions, engagement, meaning, and life satisfaction. Those who clicked on the link for a research study on positive psychology exercises were informed that they would be randomly assigned to conditions receiving either 2, 4, or 6 positive psychology exercises or an assessment-only control condition over a 6-week period and would be asked to return to the website each week to complete assessment measures. Participants completed consent online by reading the informed consent document and providing an electronic signature.

After completing the baseline questionnaires that assessed demographics as well as mood and depressive symptoms, participants were randomly assigned to one of the four groups. Those assigned to one of the three treatment conditions received instructions for the first positive psychology exercise and were asked to practice that exercise over the course of the following week. Exercises were provided in the following fixed order: three good things, strengths, gratitude visit, savoring, active-constructive responding, and life summary. Participants in the 2-exercise condition, for example, received the first two exercises in this sequence (three good things and strengths) and were encouraged to use these exercises in the following weeks and return to the website to complete follow-up measures. Participants in the assessment-only control condition simply received a message to return the following week to complete additional assessments. Participants in all groups received email reminders each week to return to the site and complete the weekly assessments. After completion of these measures, participants received the next exercise in the sequence if allowed in their study condition.

Measures

At baseline, participants completed thorough demographics that included age, gender, ethnicity, marital status, educational attainment, income, and region (first three digits of their zip code). Participants also responded to questions regarding history of psychiatric illness and treatment. During the baseline assessment and at each weekly follow-up, participants also completed the Center for Epidemiological Studies-Depression scale (CES-D), a 20-item self-report scale designed to measure the current level of depressive symptoms [33]. The CES-D contains items assessing negative symptoms such as depressed mood (eg, “I felt depressed”) and appetite disturbance (eg, “I

did not feel like eating; my appetite was poor”) as well as positive symptoms (eg, “I felt hopeful about the future”) which are reverse-coded to create a single summary scale of depressive symptoms.

Data Analyses

In order to investigate the hypotheses of the current study, we conducted chi-square tests comparing attrition rates across the four conditions. Given the high rates of attrition in Internet-based studies, we conducted analyses of the *completers*, that is, participants who completed the pretest and posttest measures. We used repeated measures analysis of covariance of the completers to assess changes over the 6-week intervention period in depressive symptoms across the 4 conditions while controlling for baseline levels of symptoms.

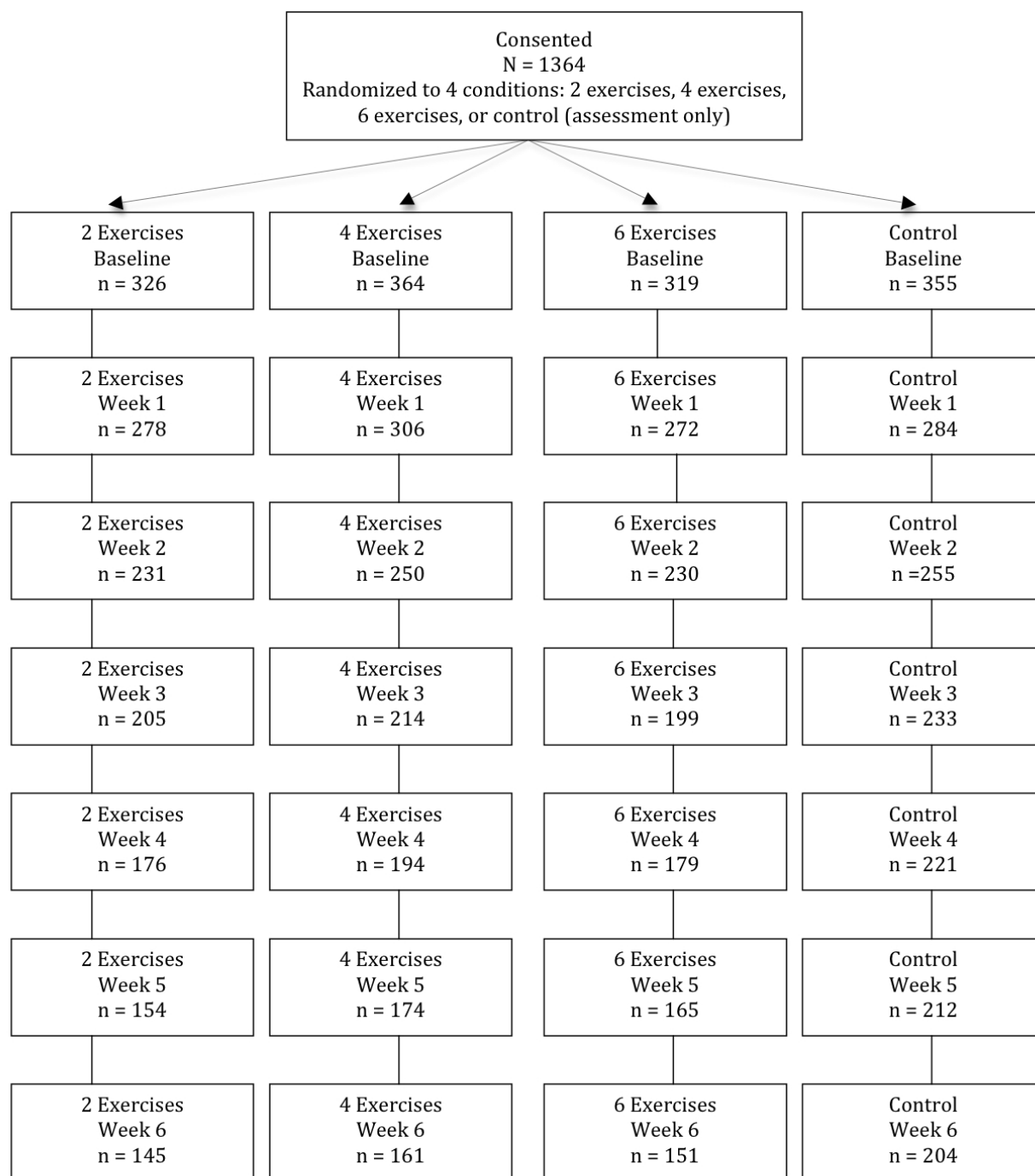
Results

Sample

Participants included all those who visited the website and consented to participate during a 20-month period from February 2007 through November 2008. Participants were predominantly women (1038/1364, 76.1%) and middle aged with a mean age of 42.38 years (SD 12.08). The largest proportion of participants lived in the United States (743/1364, 54.5%), although many participants were from other countries (621/1364, 45.5%). Most participants were married (602/1364, 44.1%), while 35.0% (477/1364) reported never having been married, and the remaining 20.9% (285/1364) were separated, divorced, or widowed. This was a highly educated sample with 77.2% (1053/1364) of participants reporting they had received a bachelor's degree or higher. Income was generally high as 24.7% (337/1364) reported earning over US \$99,000, and only 11.7% (160/1364) reported earning under US \$20,000.

Attrition

Figure 1 displays the consolidated standards of reporting trials (CONSORT) diagram for the study. Of the 1364 participants who consented to participate, 1140 (83.6%), 966 (70.8%), 851 (62.4%), 770 (56.4%), 705 (51.7%), and 661 (48.4%) completed measures each week at the week 1 through week 6 assessments respectively. At the end of the 6-week assessment period, attrition rates across the active treatment conditions were similar, that is, 55.5% (181/326) for 2 exercises, 55.8% (203/364) for 4 exercises, and 52.7% (168/319) for 6 exercises, but all were significantly greater than the 42.5% (151/355) attrition rate for the control condition ($\chi^2_3 = 16.40$, $P < .001$).

Figure 1. Flow of participants through the trial.

Use of Exercises

Compliance (ie, exercise usage) was assessed for each exercise at the 1-week assessment following its assignment and again at each of the subsequent assessments. There were no significant

differences between the 6 exercises on use of the techniques. Individuals in the 6-exercise condition were significant more likely than participants in the 4-exercise condition to use both the third ($F_{1,312} = 5.61$, $P = .02$) and fourth ($F_{1,313} = 6.03$, $P = .02$) exercises. At the end of the 6 weeks, the three active

conditions differed significantly on the number of days that the techniques were used ($F_{2,358} = 33.84$, $P < .001$). For the 2-exercise condition, mean usage was 6.1 days (SD 3.56); for the 4-exercise condition, mean usage was 7.94 days (SD 4.77); and for the 6-exercise condition, mean usage was 12.1 days (SD 7.36). For 5 of the 6 exercises (all of the 6 with exception of the life summary), enjoyment was related to continued use of the exercise at the 6-week follow-up (r 's = .12 to .39, P 's < .05).

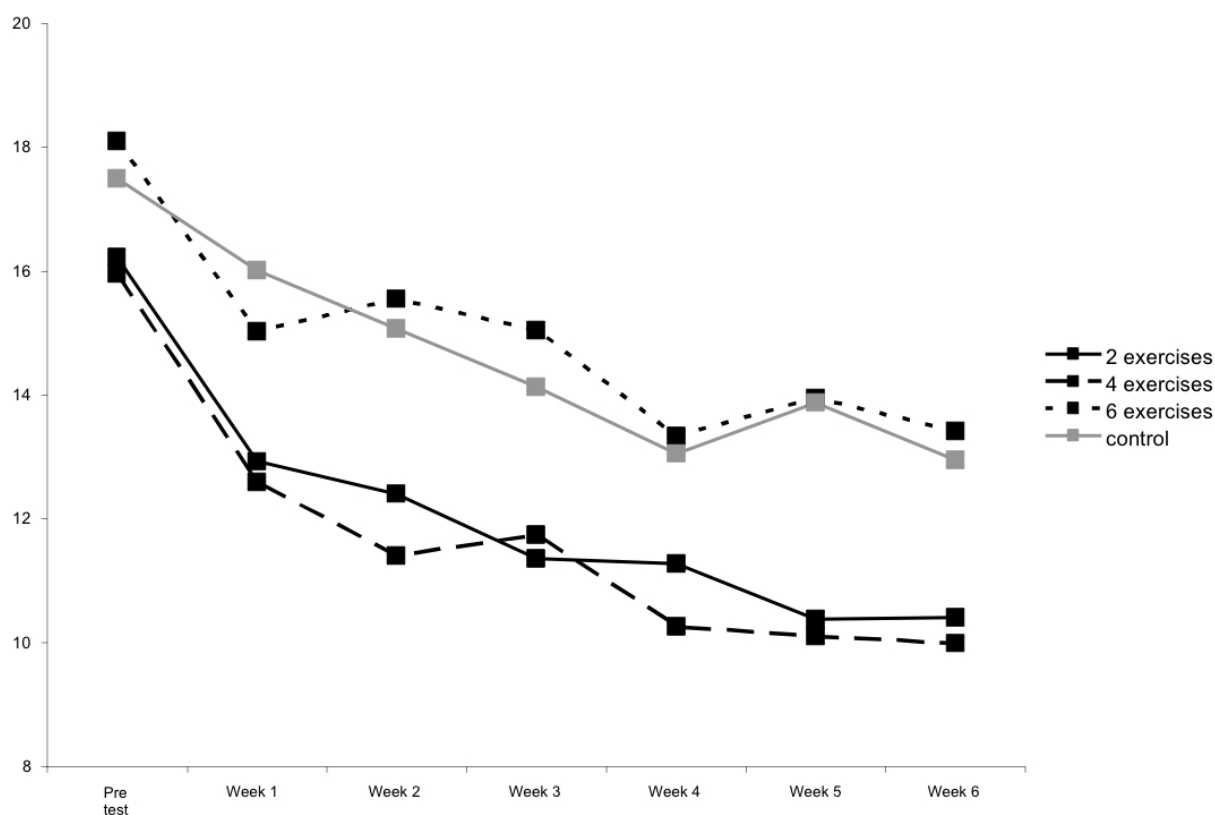
Efficacy of Exercises

To analyze efficacy of the package of interventions, a repeated measures analysis of covariance was conducted on measures of depressive symptoms. At baseline, the average scores on the CES-D were consistent with mild-to-moderate levels of depressive symptoms (mean 16.93, SD 12.64). There were no significant differences between the groups on baseline levels of depressive symptoms: means were 16.22, 15.95, and 18.10 for participants in the 2-, 4-, and 6-exercise conditions respectively, while the mean for those in the control condition was 17.49. A significant main effect of time ($F_{1,656} = 94.71$, $P < .001$) demonstrates that depression scores dropped over time in all conditions. Furthermore, a significant condition by time

interaction ($F_{3,656} = 4.77$, $P = .003$) indicates that these decreases differed across the conditions. Post-hoc Tukey HSD (honestly significant difference) tests analyzed the differences among the four conditions. These analyses showed that at the end of the intervention period, participants in both the 2-exercise condition and the 4-exercise condition reported significantly fewer depressive symptoms than those in the 6-exercise condition and numerically (but not significantly) lower than participants in the control condition. Depressive symptoms reported by participants in the 6-exercise condition did not differ significantly from those reported by participants in the control condition. Figure 2 displays mean CES-D scores by condition at each weekly assessment. Estimated marginal means in the CES-D were 12.31, 11.95, 15.89, and 14.53 for the 2-, 4-, and 6-exercise groups and control condition, respectively.

We also analyzed whether use of the exercises played a mediating role in the changes in depressive symptoms. We conducted regressions predicting changes in depressive symptoms while controlling for baseline symptoms investigating both linear and curvilinear relationships with the number of days that exercises were practiced. Neither term was significant, suggesting that the number of days of engagement did not relate to changes in depressive symptoms.

Figure 2. Mean CES-D scores by condition.



Discussion

The results of this study highlighted considerations regarding the implications of providing additional content to participants in Internet-based interventions. First, increasing the amount and variety of content did not lead to differences in attrition rates. Participants in the 2-, 4-, and 6-exercise conditions dropped out at roughly the same rate. Second, receiving more exercises corresponded to higher use of these exercises. Participants who received 6 exercises in total reported greater use of the gratitude visit and savoring exercises than those participants who only received 4 exercises. Furthermore, participants in the 6-exercise condition reported increased use summing across all the techniques at follow-up. Participants in the 6-exercise condition did report decreases in depressive symptoms after completing the interventions; however, they did not differ in reduction of depressive symptoms compared with those in the control group. Participants who received 2 or 4 exercises, on the other hand, both experienced significant reductions compared with participants in the control group and those in the 6-exercise condition. This suggests a curvilinear relationship such that some exercises are better than none but that 6 exercises was associated with fewer gains.

The main aim of this study was to translate a previously validated package of positive psychology exercises for online dissemination. Indeed, in this study, participants who, on average, experienced mild-to-moderate depressive symptoms at baseline did experience reductions in depressive symptoms. A further aim of this study was to evaluate a dose-response relationship to determine if receiving more elements of the intervention led to differences in adherence or efficacy. Increasing the content delivered to participants did not reduce rates of follow-up and corresponded to increases in reported rates of usage of the strategies taught through the intervention. This result might be because increasing the number of choices given to someone will increase the likelihood that he or she will find an exercise (or several exercises) that is a good fit and thus continue to use the technique.

Despite other research that has demonstrated a relationship between adherence to the intervention and efficacy [34,35], this study found that participants who received less content actually made larger gains than participants who received more content. This did not support our hypothesis that providing more exercises would lead to increased benefits. Instead, the results demonstrate a more complex relationship between the amount of content delivered and efficacy of the set of techniques. It might be that increasing the diversity of exercises leads to participants splitting their time among the techniques and not focusing on any of the techniques long enough to benefit substantially.

This finding, if replicated, provides the most constructive advice for those interested in creating and disseminating Internet-based interventions: do not overwhelm the participants. Past research in positive psychology interventions has found that variety is an important moderator of intervention gains and that several

techniques are more effective than a single technique [23,36]. In one such study, participants who were allowed to vary acts of kindness from a previously specified list over a 10-week period received greater gains in happiness compared with participants who were required to repeat the same acts of kindness each week [36]. Consistent with the results of the current study, however, participants who were allowed complete freedom to create new acts each week actually did worse than those who varied the acts but used the previously generated list.

These findings taken together indicate that increasing content might provide additional skills and tools to participants up to a point. But after this point, participants might not know which tools to use when or might become overwhelmed by the number of choices afforded to them. Indeed, the vast literature on choice suggests that although multiple options are beneficial, the relationship between choice and human happiness is curvilinear with too many options actually leading to various psychological and economic factors (time trade-offs, remorse, rumination, etc) that decrease psychological well-being [37]. Although the issue of depth versus breath of options is still an open and important empirical consideration, it appears that in some cases at least less is more.

The findings of this study are tempered by its limitations. Similar to other Internet-based studies, participants in this study were overwhelming female, highly affluent, and well educated. Although this study was an open trial and participants throughout the world could (and did) visit the website, the sample was not representative of the world population. In part, this might be due to the recruitment procedure, which advertised via a website associated with a popular positive psychology book and the University of Pennsylvania's Positive Psychology Center. Future research on these interventions should adopt alternative recruitment procedures to draw people into the study rather than relying on participant flow from existing resources. Another limitation is that although this study did experimentally manipulate the amount of content participants were offered, the content was presented through an automated system, and thus we cannot be sure whether participants actually read and understood the instructions. Furthermore, instructions were delivered over subsequent weeks, and participants were asked to use one activity per week although in the real world, participants are more likely to receive several potential activities at once and use them at will. It would be worthwhile to further investigate the question of depth versus breath of interventions in a format where participants can access the additional content simultaneously.

Overall, this study provides a further illustration of the flexibility and adaptability of positive psychology interventions. As a majority of the existing research on Internet-based interventions has evaluated cognitive-behavioral based techniques, future work should continue to integrate other validated strategies to promote psychological well-being. The science of Internet interventions can be advanced through expanding options and strategies to promote worldwide well-being.

Acknowledgments

Stephen Schueller is now at the Department of Psychiatry, University of California, San Francisco. Acacia Parks is now at Hiram College. This work was completed at the Positive Psychology Center at the University of Pennsylvania. We are especially grateful to Marty Seligman for his support and guidance in completing this research and manuscript. We would also like to recognize Patty Newbold and Peter Schulman for their contributions to creating and hosting the technical aspects of the Web intervention and Ricardo Muñoz for helpful comments on an early version of this manuscript.

Conflicts of Interest

None declared.

References

1. Mental Health. Geneva, Switzerland: World Health Organization; 2011. Depression URL: http://www.who.int/mental_health/management/depression/definition/en/ [accessed 2011-05-13] [WebCite Cache ID 5yeamj3CT]
2. Murray CJ, Lopez AD. Evidence-based health policy--lessons from the Global Burden of Disease Study. *Science* 1996 Nov 1;274(5288):740-743. [Medline: [8966556](#)]
3. Lewinsohn PM, Clarke GN, Seeley JR, Rohde P. Major depression in community adolescents: age at onset, episode duration, and time to recurrence. *J Am Acad Child Adolesc Psychiatry* 1994;33(6):809-818. [doi: [10.1097/00004583-199407000-00006](#)] [Medline: [7598758](#)]
4. Flett GL, Vredenburg K, Krames L. The continuity of depression in clinical and nonclinical samples. *Psychol Bull* 1997 May;121(3):395-416. [Medline: [9136642](#)]
5. Judd LL, Akiskal HS, Paulus MP. The role and clinical significance of subsyndromal depressive symptoms (SSD) in unipolar major depressive disorder. *J Affect Disord* 1997 Aug;45(1-2):5-17; discussion 17. [Medline: [9268771](#)]
6. Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatr Scand* 2004 May;109(5):325-331. [doi: [10.1111/j.1600-0447.2004.00301.x](#)] [Medline: [15049768](#)]
7. Seligman ME, Csikszentmihalyi M. Positive psychology. An introduction. *Am Psychol* 2000 Jan;55(1):5-14. [Medline: [11392865](#)]
8. Seligman ME. *Flourish: A Visionary New Understanding of Happiness and Well-being*. New York, NY: Free Press; 2011.
9. Lyubomirsky S, King L, Diener E. The benefits of frequent positive affect: does happiness lead to success? *Psychol Bull* 2005 Nov;131(6):803-855. [doi: [10.1037/0033-2909.131.6.803](#)] [Medline: [16351326](#)]
10. Algoe SB, Haidt J, Gable SL. Beyond reciprocity: gratitude and relationships in everyday life. *Emotion* 2008 Jun;8(3):425-429. [doi: [10.1037/1528-3542.8.3.425](#)] [Medline: [18540759](#)]
11. Schueller SM, Seligman ME. Pursuit of pleasure, engagement, and meaning: Relationships to subjective and objective measures of well-being. *The Journal of Positive Psychology* 2010;5(4):253-263. [doi: [10.1080/17439761003794130](#)]
12. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. *PLoS Med* 2010 Jul;7(7):e1000316 [FREE Full text] [doi: [10.1371/journal.pmed.1000316](#)] [Medline: [20668659](#)]
13. Tugade MM, Fredrickson BL. Resilient individuals use positive emotions to bounce back from negative emotional experiences. *J Pers Soc Psychol* 2004 Feb;86(2):320-333. [doi: [10.1037/0022-3514.86.2.320](#)] [Medline: [14769087](#)]
14. Fredrickson BL, Tugade MM, Waugh CE, Larkin GR. What good are positive emotions in crises? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. *J Pers Soc Psychol* 2003 Feb;84(2):365-376. [Medline: [12585810](#)]
15. Wood AM, Joseph S. The absence of positive psychological (eudemonic) well-being as a risk factor for depression: a ten year cohort study. *J Affect Disord* 2010 May;122(3):213-217. [doi: [10.1016/j.jad.2009.06.032](#)] [Medline: [19706357](#)]
16. Keyes CL. The mental health continuum: from languishing to flourishing in life. *J Health Soc Behav* 2002 Jun;43(2):207-222. [Medline: [12096700](#)]
17. Fordyce MW. Development of a program to increase personal happiness. *Journal of Counseling Psychology* 1977;24(6):511-521. [doi: [10.1037/0022-0167.24.6.511](#)]
18. Emmons RA, McCullough ME. Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life. *J Pers Soc Psychol* 2003 Feb;84(2):377-389. [Medline: [12585811](#)]
19. Lyubomirsky S, Sheldon KM, Schkade D. Pursuing happiness: The architecture of sustainable change. *Review of General Psychology* 2005;9(2):111-131. [doi: [10.1037/1089-2680.9.2.111](#)]
20. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. *Am Psychol* 2005;60(5):410-421. [doi: [10.1037/0003-066X.60.5.410](#)] [Medline: [16045394](#)]
21. Peterson C, Seligman ME. *Character strengths and virtues: a handbook and classification*. Oxford: Oxford Univ Pr, 2004.; 2004.
22. Bryant FB, Smart CM, King SP. Using the past to enhance the present: Boosting happiness through positive reminiscence. *J Happiness Stud* 2005;6(3):227-260. [doi: [10.1007/s10902-005-3889-4](#)]

23. Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. *J Clin Psychol* 2009 May;65(5):467-487. [doi: [10.1002/jclp.20593](https://doi.org/10.1002/jclp.20593)] [Medline: [19301241](https://pubmed.ncbi.nlm.nih.gov/19301241/)]
24. Seligman ME, Rashid T, Parks AC. Positive psychotherapy. *Am Psychol* 2006 Nov;61(8):774-788. [doi: [10.1037/0003-066X.61.8.774](https://doi.org/10.1037/0003-066X.61.8.774)] [Medline: [17115810](https://pubmed.ncbi.nlm.nih.gov/17115810/)]
25. Seligman ME, Ernst RM, Gillham J, Reivich K, Linkins M. Positive education: Positive psychology and classroom interventions. *Oxford Review of Education* 2009;35(3):293-311. [doi: [10.1080/03054980902934563](https://doi.org/10.1080/03054980902934563)]
26. Reivich KJ, Seligman ME, McBride S. Master resilience training in the U.S. Army. *Am Psychol* 2011 Jan;66(1):25-34. [doi: [10.1037/a0021897](https://doi.org/10.1037/a0021897)] [Medline: [21219045](https://pubmed.ncbi.nlm.nih.gov/21219045/)]
27. Brunwasser SM, Gillham JE, Kim ES. A meta-analytic review of the Penn Resiliency Program's effect on depressive symptoms. *J Consult Clin Psychol* 2009 Dec;77(6):1042-1054. [doi: [10.1037/a0017671](https://doi.org/10.1037/a0017671)] [Medline: [19968381](https://pubmed.ncbi.nlm.nih.gov/19968381/)]
28. Seligman ME. *Authentic Happiness: Using the New Positive Psychology to Realize Your Potential for Lasting Fulfillment*. New York: Free Press; 2002.
29. Munson S, Lauterbach D, Newman M, Resnick P. Happier Together: Integrating a Wellness Application Into a Social Network Site. *Persuasive* 2010:27-39. [doi: [10.1007/978-3-642-13226-1_5](https://doi.org/10.1007/978-3-642-13226-1_5)]
30. Parks AC, Della Porta MD, Pierce RS, Zilca R, Lyubomirsky, S. How do people pursue happiness in their everyday lives?: The characteristics and behaviors of happiness seekers. *Emotion*, in press 2012 (forthcoming).
31. Geraghty AW, Wood AM, Hyland ME. Attrition from self-directed interventions: investigating the relationship between psychological predictors, intervention content and dropout from a body dissatisfaction intervention. *Soc Sci Med* 2010 Jul;71(1):30-37. [doi: [10.1016/j.socscimed.2010.03.007](https://doi.org/10.1016/j.socscimed.2010.03.007)] [Medline: [20400220](https://pubmed.ncbi.nlm.nih.gov/20400220/)]
32. Gable SL, Reis HT, Impett EA, Asher ER. What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *J Pers Soc Psychol* 2004 Aug;87(2):228-245. [doi: [10.1037/0022-3514.87.2.228](https://doi.org/10.1037/0022-3514.87.2.228)] [Medline: [15301629](https://pubmed.ncbi.nlm.nih.gov/15301629/)]
33. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement* 1977;1(3):385-401. [doi: [10.1177/014662167700100306](https://doi.org/10.1177/014662167700100306)]
34. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. *Nicotine Tob Res* 2009 Sep;11(9):1025-1034. [doi: [10.1093/ntr/ntp090](https://doi.org/10.1093/ntr/ntp090)] [Medline: [19640833](https://pubmed.ncbi.nlm.nih.gov/19640833/)]
35. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ* 2004 Jan 31;328(7434):265 [FREE Full text] [doi: [10.1136/bmj.37945.566632.EE](https://doi.org/10.1136/bmj.37945.566632.EE)] [Medline: [14742346](https://pubmed.ncbi.nlm.nih.gov/14742346/)]
36. Tkach CT. *Unlocking the treasury of human kindness: Enduring improvements in mood, happiness, and self-evaluations*. Dissertation Abstracts International: Section B: The Sciences and Engineering 2006.
37. Schwartz B, Ward A, Monterosso J, Lyubomirsky S, White K, Lehman DR. Maximizing versus satisficing: happiness is a matter of choice. *J Pers Soc Psychol* 2002 Nov;83(5):1178-1197. [Medline: [12416921](https://pubmed.ncbi.nlm.nih.gov/12416921/)]

Abbreviations

CES-D: Center for Epidemiological Studies–Depression scale
CONSORT: consolidated standards of reporting trials
HSD: honestly significant difference
PPT: positive psychotherapy

Edited by G Andersson, H Christensen, A Calear; submitted 24.05.11; peer-reviewed by A Wood, R Johansson; comments to author 16.06.11; revised version received 27.07.11; accepted 29.07.11; published 25.06.12.

Please cite as:

Schueller SM, Parks AC

Disseminating Self-Help: Positive Psychology Exercises in an Online Trial

J Med Internet Res 2012;14(3):e63

URL: <http://www.jmir.org/2012/3/e63/>

doi: [10.2196/jmir.1850](https://doi.org/10.2196/jmir.1850)

PMID: [22732765](https://pubmed.ncbi.nlm.nih.gov/22732765/)

©Stephen M. Schueller, Acacia C. Parks. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 25.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic

information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Self-monitoring Using Mobile Phones in the Early Stages of Adolescent Depression: Randomized Controlled Trial

Sylvia Deidre Kauer^{1,2}, BBehavSci (Hons); Sophie Caroline Reid¹, MPsyCh(Clinical), PhD; Alexander Hew Dale Crooke¹, BSocSc(Hons); Angela Khor¹, BA(Hons); Stephen John Charles Hearps¹, BPsych, PGDipPsych; Anthony Francis Jorm³, PhD, DSc; Lena Sancic⁴, MBBS, PhD, FRACGP; George Patton^{1,5}, MD, FRANZCP

¹Centre for Adolescent Health, Royal Children's Hospital & Murdoch Childrens Research Institute, Parkville, VIC, Australia

²School of Behavioural Science, University of Melbourne, Melbourne, VIC, Australia

³Orygen Youth Health Research Centre, Centre for Youth Mental Health, University of Melbourne, Parkville, VIC, Australia

⁴Department of General Practice, University of Melbourne, Carlton, VIC, Australia

⁵Department of Paediatrics, University of Melbourne, Melbourne, Australia

Corresponding Author:

Sylvia Deidre Kauer, BBehavSci (Hons)

Centre for Adolescent Health

Royal Children's Hospital & Murdoch Childrens Research Institute

Flemington Rd

Parkville, VIC, 3052

Australia

Phone: 61 393455890

Fax: 61 393456502

Email: sylvia.kauer@mcri.edu.au

Abstract

Background: The stepped-care approach, where people with early symptoms of depression are stepped up from low-intensity interventions to higher-level interventions as needed, has the potential to assist many people with mild depressive symptoms. Self-monitoring techniques assist people to understand their mental health symptoms by increasing their emotional self-awareness (ESA) and can be easily distributed on mobile phones at low cost. Increasing ESA is an important first step in psychotherapy and has the potential to intervene before mild depressive symptoms progress to major depressive disorder. In this secondary analysis we examined a mobile phone self-monitoring tool used by young people experiencing mild or more depressive symptoms to investigate the relationships between self-monitoring, ESA, and depression.

Objectives: We tested two main hypotheses: (1) people who monitored their mood, stress, and coping strategies would have increased ESA from pretest to 6-week follow-up compared with an attention comparison group, and (2) an increase in ESA would predict a decrease in depressive symptoms.

Methods: We recruited patients aged 14 to 24 years from rural and metropolitan general practices. Eligible participants were identified as having mild or more mental health concerns by their general practitioner. Participants were randomly assigned to either the intervention group (where mood, stress, and daily activities were monitored) or the attention comparison group (where only daily activities were monitored), and both groups self-monitored for 2 to 4 weeks. Randomization was carried out electronically via random seed generation, by an in-house computer programmer; therefore, general practitioners, participants, and researchers were blinded to group allocation at randomization. Participants completed pretest, posttest, and 6-week follow-up measures of the Depression Anxiety Stress Scale and the ESA Scale. We estimated a parallel process latent growth curve model (LGCM) using Mplus to test the indirect effect of the intervention on depressive symptoms via the mediator ESA, and calculated 95% bias-corrected bootstrapping confidence intervals (CIs).

Results: Of the 163 participants assessed for eligibility, 118 were randomly assigned and 114 were included in analyses (68 in the intervention group and 46 in the comparison group). A parallel process LGCM estimated the indirect effect of the intervention on depressive symptoms via ESA and was shown to be statistically significant based on the 95% bias-corrected bootstrapping CIs not containing zero (−6.366 to −0.029). The proportion of the maximum possible indirect effect estimated was $\kappa^2 = .54$ (95% CI .426–.640).

Conclusions: This study supported the hypothesis that self-monitoring increases ESA, which in turn decreases depressive symptoms for young people with mild or more depressive symptoms. Mobile phone self-monitoring programs are ideally suited to first-step intervention programs for depression in the stepped-care approach, particularly when ESA is targeted as a mediating factor.

Trial Registration: ClinicalTrials.gov NCT00794222; <http://clinicaltrials.gov/ct2/show/NCT00794222> (Archived by WebCite at <http://www.webcitation.org/65lldW34k>)

(*J Med Internet Res* 2012;14(3):e67) doi:[10.2196/jmir.1858](https://doi.org/10.2196/jmir.1858)

KEYWORDS

Mobile phone; early intervention; patient monitoring; randomized controlled trial; consciousness; depressive disorder; affect

Introduction

Depression is a common, recurrent disorder and contributes to a substantial burden of disease [1]. Up to 25% of the population experience at least one episode of depression in their lifetime [2] with more women affected by depression than men [2,3]. First episodes of depression generally begin during adolescence [4] with up to 30% of young people experiencing mild subclinical depressive symptoms by 18 years of age [5,6] and many progressing to major depressive disorder [6]. A meta-analysis of the depressive disorder point prevalence estimates for 13- to 18-year-olds reported that 5.7% of adolescents may have a depressive disorder diagnosed at any one point in time [7]. The severity and longevity of depression are associated with substantial economic costs [8], attributed to the early onset and recurrent nature of depression. The World Health Organization predicted that depression will be the second leading cause of disease burden worldwide by 2020 [1], indicating the need for new methods to reduce the burden of depression. Current interventions are estimated to cover 66% of the burden of depression, with 33% of this burden unable to be averted by available treatments [4]. The large number of people who experience mild depressive symptoms account for more burden of depression than those with major depressive disorder [9], suggesting that new methods are needed and that they should focus on the early stages of depression.

The stepped-care approach, beginning with simple, inexpensive interventions (that are given before the onset of diagnosed major depressive disorder) with the ability to step up to higher-intensity programs as needed [9-11], has the advantage of providing a low-intensity early intervention with reduced length and cost of the treatment to young people experiencing mild mental health symptoms [12,13]. Early interventions that focus on intensive computerized or school-based cognitive behavioral therapy-based programs may be best regarded as second-step interventions, as they are time consuming and have high attrition rates [14]. Third-step interventions may include individualized face-to-face therapy, whereas a combination of face-to-face therapy and antidepressants should be considered only as a last resort for teenagers [11]. Van Straten et al [11] advocate the use of watchful waiting as a first step in the stepped-care model. The watchful waiting approach has two limitations. First, as general practitioners (GPs) are often the first point of contact for people with mild mental health symptoms, the onus of watchful waiting would be placed on GPs. GPs, however, are under pressure to treat many people within a day and to keep

appointment times brief (the current average appointment time is 15 minutes) [15]. Second, a young person's lack of awareness of emotional distress may reduce the effectiveness of the watchful waiting approach, as GPs are more likely to identify mental health problems in young people who are aware of emotion distress [16]. Short programs involving the completion of homework diaries have been shown to have larger effects on depressive symptoms than longer-duration programs do and offer a potential alternative first-step approach to watchful waiting [17]. Such programs may be completed via mobile phones as a first-step early intervention with the opportunity to step up to more intensive treatment programs as needed.

Mobile Phone Self-monitoring Program

Methods for stepped-care interventions involving technology, such as computers, the Internet, or mobile phones, and simple methods such as self-monitoring techniques can engage young people and foster their involvement [18-20]. Technology is particularly important when targeting young people, dubbed early adopters, as they are familiar with and rely on technology for much of their social interactions with peers and readily engage with electronic devices [18-20]. In particular, studies conducted using electronic self-monitoring devices with young people have demonstrated that young people readily engage with the technology [21-23] and have high compliance [24].

We developed and piloted a mobile phone self-monitoring program, the mobiletype program (Mobile Tracking of Young People's Experiences) [21,23,25], well suited as a first step in the stepped-care approach. In our pilot studies, young people were asked to monitor their mood, stress, and daily activities on a mobile phone application 4 times a day for 1 week. Results suggested that young people complied with self-monitoring of mental health symptoms for the purpose of reviewing these data with their GP and found it a simple, easy-to-use tool to track their daily experiences. Both GPs and patients found communicating information via the mobiletype program to be beneficial, and the program enabled GPs to better understand the young person's mental health. The self-monitoring data can be uploaded to GPs and used to detect depressive symptoms and other mental health information to allow progression to further higher-intensity intervention if needed. Self-monitoring approaches are typically used in psychotherapy as an adjunct to cognitive behavioral therapy, with research demonstrating that self-monitoring increases the benefits of therapy more than therapy alone [26]. Research focusing on momentary sampling techniques demonstrates that self-monitoring itself can lead to

a change in behavior and that changes are generally made in a favorable direction [27-29].

Mechanism Underlying Self-monitoring

The contents of early intervention programs generally target a mechanism that predicts the outcome [30]. For example, cognitive-behavioral-based universal programs focus on cognitive restructuring and problem-solving skills training [31], which are predicted to reduce depressive symptoms. In this study, the intervention was based on emotional self-awareness (ESA), which is hypothesized to predict depressive symptomatology. Increasing awareness of emotions is an important therapeutic step in most psychotherapies for depression and other mental illnesses [32] by preparing individuals for changing their cognitions, beliefs, and schemas [33]. A recent case study using mobile phones [34] showed that self-monitoring increased positive mood and coping strategies while decreasing negative mood in four adult employees with stress. Participants' ESA increased, and they were able to internalize the questions and therapies used in the mobile program on completion of the study. Self-monitoring studies are promising; however, further research is needed to explore the possible mental health benefits of self-monitoring techniques with young people who have symptoms of depression.

ESA may also provide a suitable framework for first-step intervention programs by assisting young people to become aware of their emotions in preparation for learning more adaptive coping strategies. Adolescence is an ideal target for first-step intervention of mental health problems, as young people begin to develop the ability to independently cope with everyday stresses and negative emotions during their teen years [35]. The current study focused on examining the relationship between self-monitoring and depressive symptoms through the mechanism of ESA, specifically focusing on (1) recognizing emotions [32,36,37], (2) being able to identify emotions [38-40], (3) identifying contextual factors surrounding emotions [36,38,41], (4) communicating emotional states and associated factors to others and internally [42-44], and (5) planning and making decisions about how to cope with an emotional state [38,45,46]. Through self-monitoring, young people can learn to recognize emotional states, and to identify and differentiate various emotions within different contexts, leading to effective communication of emotions to others and productive decision making.

The overall aim of the current study was to investigate, in a randomized controlled trial, the utility of the mobiletype program as a first-step intervention program. The primary hypothesis was that young people who completed the mobiletype intervention program would have lower depressive symptoms than those who completed the attention comparison program. Using a mixed-methods model, we found that depressive symptoms significantly decreased over time for both intervention and attention comparison groups, and we found no significant difference in mental health symptoms between groups [47]. We attributed the lack of difference between groups in mental health symptoms to the unanticipated effect of the training, resources, and support given to young people and GPs over the course of the study for both the intervention and the comparison groups.

Interestingly, the intervention group was found to have a significant increase in ESA compared with the comparison group. As both groups were found to have significantly decreased mental health symptoms, there was reduced power to detect a difference between groups.

The goal of this secondary analysis was to further examine the effects of a mobile phone self-monitoring program on ESA among young people with mild or more depressive symptoms as a first-step treatment of depression. Specifically, we were interested in the effects of self-monitoring on ESA and the association between ESA and depressive symptoms. We hypothesized that self-monitoring mood, stress, and coping strategies would increase young people's awareness of their emotions, which would lead to a decrease in their symptoms of depression. We tested the following hypotheses: (1) young people in the intervention group would have an increase in ESA from pretest to 6-week follow-up compared with the attention comparison group, and (2) the increase in ESA would predict a decrease in depressive symptoms. We estimated a parallel process latent growth curve model (LGCM) [30,48] to investigate whether ESA would mediate the relationship between the intervention program and depressive symptoms. A secondary analysis compared the effects of the mobiletype intervention and comparison programs on rumination, with the hypothesis that the intervention group would have decreased rumination compared with the comparison group.

Methods

Trial Design

The data presented here were from the mobiletype randomized controlled trial conducted from 2009 to 2011. This was a multicenter, multiregional, stratified (according to region), single-blind, attention-controlled study with balanced (1:1) individual randomization into parallel groups. This study was conducted in Victoria, Australia adhering to the reporting recommendations from the CONSORT statement [49] as a guide.

Participants

General Practitioners

All GPs in the Goulburn Valley Region and Albury-Wodonga Region were invited to participate in the study via the Regional Division of General Practice (support units that service clinical practices within a region). GPs in Melbourne were recruited via the local Divisions of General Practice. Clinics were targeted that listed an interest in adolescent health on the Melbourne General Practice Network (www.mgpn.com.au). Participating GPs were trained to use the mobiletype website and were provided with a study manual that included the study procedure, a variety of clinical supports (including referral details of adolescent-friendly allied health professionals and services), youth-friendly Internet, email, and phone support, and youth-focused psychoeducation handouts on a range of mental health problems (this information was also available on the mobiletype website), which were available for all participating GPs and patients. Continuing professional development quality assurance points were available to GPs for their participation

in the study. Of the 103 GPs who agreed to participate, 35 actively participated in the study with at least one young person. These contributing GPs were from 26 different practices in the three recruitment areas: 12 in greater Melbourne, 7 in Albury-Wodonga, and 7 in the Goulburn Valley.

Young People

Young people meeting the following inclusion criteria were eligible to participate regardless of their reason for visiting the GPs. Participants were required to (1) be aged between 14 and 24 years, (2) speak proficient English, and (3) have a mild or more severe emotional or mental health issue as assessed by their GP or indicated by a score greater than 16 on the Kessler Psychological Distress Scale [50]. Participants were excluded if their psychiatric or medical condition prevented them from complying with either the requirements of informed consent or the study protocol (ie, current psychosis or imminent hospitalization).

Mobiletype Program

We used Version 4 of the mobiletype program as the intervention and attention comparison in this study, which was created in-house using Java Platform, Micro Edition by the Murdoch Childrens Research Institute. This program was written for use with multiple models of mobile phones and firmware. For this trial, participants were lent a study mobile phone with either the mobiletype intervention or a comparison program

uploaded onto it. Data from the program was uploaded to a secure website constructed and hosted by the Murdoch Childrens Research Institute as well as encrypted and stored on the mobile phones.

Participants were prompted to complete a mobiletype entry by an auditory signal (beep) emitted from the mobile phone at random intervals in the blocks outlined in Table 1. If no report was completed the phone emitted one reminder signal after 5 minutes. Entries were time coded and saved. Participants were also able to complete the program any time and were able to complete an entry between 10 PM and 8 AM, although no trigger was sent at this time. Entries from 10 PM to 12 AM consisted of the evening questions and entries from 12 to 8 AM consisted of the same questions as the afternoon questions, as shown in Table 1. Each report took approximately 1–3 minutes to complete.

Intervention Program

The intervention group monitored themselves using the complete mobiletype program, which assessed 8 areas of functioning as developed in previous mobiletype studies [21,23], consisting of current activities, location, companions, mood, recent stressful events, responses to stressful events, alcohol use, cannabis use, quality and quantity of sleep, quantity and type of exercise, and diet. The time of day when each module assessing the eight areas was delivered varied as displayed in Table 1.

Table 1. Modules included in each block of the mobiletype comparison and intervention programs.

Module	Morning 8–10:59 AM	Noon 11 AM to 3:29 PM	Afternoon 3:30–7:59 PM	Evening 8–10 PM
Intervention program				
Current activity	✓	✓	✓	✓
Stress	✓	✓	✓	✓
Mood	✓	✓	✓	✓
Alcohol use		✓		
Cannabis use		✓		
Sleep	✓			
Diet				✓
Exercise				✓
Comparison program				
Current activity	✓	✓	✓	✓
Stress				
Mood				
Alcohol use				
Cannabis use				
Sleep	✓			
Diet				✓
Exercise				✓

Comparison Program

The attention comparison program was designed to provide a data collection process similar to that in the intervention group by controlling for the amount of time spent engaged in the program condition and the overall research methodology and the attention given to them by health care professionals and research staff [51]. The comparison group monitored themselves using an abbreviated version of the mobiletype program, which assessed only current activities, location, companions, quality and quantity of sleep, quantity and type of exercise, and diet (meals, snacks, junk food, and soft drinks consumed). Importantly, we removed the modules pertaining to ESA and mental health as shown in Table 1 (ie, mood, stress, alcohol and cannabis use).

Summary Reports

Data collected by the mobiletype program (intervention and comparison groups) on the mobile phone was sent via short message service to a secure website, where it was automatically collated. Each area of assessment was displayed in graphs (eg, daily mood graphs) or in tables (eg, daily alcohol intake). An individualized summary report of the data was written following structured prescriptive guidelines by the second author (registered psychologist), or the first author under the supervision of the second author.

Outcome Measures

The pretest, posttest, and 6-week follow-up questionnaire packages included the Depression Anxiety Stress Scale (DASS) [52] and the ESA Scale. The DASS is a 21-item response form with subscales of depression, anxiety, and stress with scores ranging from 0 to 42. The DASS has Australian norms and clinically validated ranges. A high DASS score indicates greater levels of depression, anxiety, or stress symptoms. The ESA Scale was adapted from the 20-item Self-reflection and Insight Scale [53], the 10-item Ruminative Response Scale [54], and the 12-item Meta-Evaluation Scale [46]. As there is no overall measure of ESA that covers the five areas of recognition, identification, communication, contextualization, and decision making, we adapted 33 items pertaining to these areas from the above scales (see Multimedia Appendix 1). These were then combined to create a total ESA scale with high internal consistency (Cronbach alpha = .83). The total ESA score ranged from 0 to 132, with higher scores indicating more ESA. Rumination was measured by the brooding subscale of the Ruminative Response Scale, which consisted of 5 items ranging from 0 to 4. A higher score indicates higher rumination.

Sample Size

We anticipated recruitment of 200 participants from 10 general practices. This sample size was based on Cohen's [55] statistical testing for multiple regression with two independent variables (accounting for the mediating variable and the outcome) to detect a medium effect with 80% power and a probability of a type I error of .05. We selected a medium effect size, as we considered this to be clinically significant. Using Fritz and MacKinnon's statistical tests [56], a sample size of 71 participants inclusive should be sufficient to detect a medium mediation effect with 80% power with bias-corrected

bootstrapping (see below). The anticipated sample size of 200 was not met due to delays in recruitment during school holidays and the H1N1 influenza pandemic [57]. As a result, we set a deadline for stopping recruitment and recruited a total of 118 participants.

Randomization

Participants were randomly assigned to either (1) the mobiletype monitoring intervention group or (2) the mobiletype attention comparison group; both groups also received medical care as usual. A database was set up by an in-house computer programmer with identification numbers for 100 Melbourne, 50 Goulburn Valley, and 50 Albury-Wodonga participants. Each number was attached to a link that downloaded either the intervention or comparison program directly to the mobile phone. This process was blinded; the intervention and comparison program could not be differentiated when downloading the program. The programmer used a random seed generator to allocate each program to the 200 identification numbers at the individual level and stratified according to area (Melbourne, Goulburn Valley, and Albury-Wodonga). This programmer was not involved in any data collection or analysis. A research assistant downloaded each program by selecting the next consecutive link for the next study mobile phone and was blinded to the allocation, as he knew only the identification number and area to load onto study mobile phones (eg, Melbourne01, Melbourne02). Mobile phones and identification numbers were allocated to consecutively recruited participants. The researchers, participants, and GPs were blinded to randomization pretest. GPs and participants became aware of the group allocation at the posttest when the summary reports were reviewed. This study had approval from the Human Research Ethics Committee of the Royal Children's Hospital, Melbourne (RCH HREC: 28113), and was registered in ClinicalTrials.gov (Reference: NCT00794222).

Procedure

Recruitment

In addition to providing treatment as usual, GPs screened their patients for eligibility to the study; organized an appointment for willing participants with a research assistant using an online booking form or a faxed referral form, or by phone; and completed a pretest questionnaire for each participant. Participants then met with a mobiletype research assistant, generally within 5 days of referral, to learn the study process, complete consent forms and the pretest questionnaire package, familiarize themselves with the mobiletype program and the other features of the phone, and complete a practice entry of the mobiletype program. Participants were provided with a study manual that described the research procedure and offered trouble-shooting tips.

Mobile Phone Monitoring Period

All participants borrowed a Sony Ericsson 7501i (Sony Limited Australia, North Ryde, NSW, Australia) mobile phone containing the mobiletype program for the study period. Information regarding the development and testing of the mobiletype program has been previously published [21]. Participants were requested to complete at least two mobiletype

entries a day until they returned for their medical review in 2 to 4 weeks. Participants and GPs were advised that 2 to 4 weeks' monitoring was the ideal monitoring period. Participants were given a subscriber identity module (SIM) card containing A\$30 in credit as partial reimbursement for their time and phone credit used.

Posttest Review

On completing the monitoring period, participants reviewed the self-monitoring data with their GP on the mobiletype website. Young people completed a posttest assessment immediately following this appointment, and again at 6 weeks and 6 months after this review (6-month posttests not included in the current analysis). GPs completed a posttest questionnaire immediately after the appointment. Questionnaires were completed online, over the phone with a researcher, or via a mailed hardcopy questionnaire. Participants were given a A\$20 gift card for each posttest questionnaire completed (maximum of A\$60 for all questionnaires completed).

Analyses

We conducted all analyses on an intention-to-treat basis using all available data from participants included at randomization. Data were assumed to be missing at random [58], and maximum likelihood estimation was used with missing data accounted for by the missing routines in the statistical program used.

Parallel Process Latent Growth Curve Model

Recently, structural equation modeling has modernized Baron and Kenny's well-known mediation model [59,60]. LGCM is a structural equation modeling technique that allows for examination of inter- and intrapersonal changes and, importantly, parallel process LGCM accommodates longitudinal data for situations in which both the mediator and outcome change over time [61]. Baron and Kenny's requirement of a statistically significant pathway between the independent variable and the outcome [59] is no longer recommended, as detecting a total effect between the independent variable and outcome reduces power, therefore reducing the likelihood that a mediated effect can be detected [56,62]. Parallel process LGCM based on Cheong et al [30] was used to test the hypothesis that the mobiletype program would increase ESA and that the change in ESA over time would change (decrease) depressive symptoms, by using the software program Mplus Version 6.11 (Muthén & Muthén, Los Angeles, CA, USA). The model estimates two latent variables, the intercept (estimated starting point) and the slope (estimated growth curve), using repeated measures for the mediator (ESA) and the outcome (depression) across the three time points. The latent intercept for both ESA and depression had factor loadings set to 1 to represent the starting point of the growth trajectory, and the factor loadings of the latent linear slope were fixed to 0, 3, and 9 to represent the time between tests: pretest (week 0), posttest (week 2–4), and 6-week follow-up (week 8–10), respectively. Group was defined as 0 for the comparison group and 1 for the intervention group. A combination of at least three of the following cut-off values was used to determine goodness of fit of the model as recommended [63,64]: comparative fit index > .95, Tucker-Lewis Index > .95, standardized root mean square

residual < .08, and root mean square error of approximation and root mean square error of approximation < .05. Another statistic considered was chi-square divided by degrees of freedom, with values less than 2 indicating a good fit and values up to 3 indicating an acceptable fit [64]. When specifying the model, we allowed the variances of the observed variables to covary at each time point and calculated all regression pathways. Indirect effects of group on the slope of depression via the intercept and the slope of ESA were also specified. As recommended [56,62,65], we used bootstrapping with resampling at 5000, with bias-corrected confidence intervals reported for the indirect pathway [66]. Resampling provides more accurate results with fewer assumptions about the data (such as normal distributions) [60].

Indirect Effect Sizes

A recent development in statistical analysis has led to the capacity of estimating effect sizes for mediation models [67]. Using R for Mac OS X GUI 1.40-devel, an open-source language and environment for statistical analysis (<http://www.r-project.org/>), we calculated effect sizes of the indirect pathways using the mediation function in the MBESS (<http://cran.r-project.org/web/packages/mediation/mediation.pdf>) in R. Bootstrapping was used to calculate 95% confidence intervals. As recommended [60,67], both the unstandardized indirect effect size and the proportion of the maximum possible indirect effect (κ^2) are reported and described in the results. Other estimates of effect size obtained in this analysis are listed in [Multimedia Appendix 2](#).

Rumination

We conducted a secondary analysis to determine whether the intervention group had a decrease in rumination compared with the comparison group. A mixed model analysis was conducted over time and between groups using SPSS Version 17.0.0 (IBM Corporation, Somers, NY, USA) with the MIXED procedure. As with the LGCM, survey time was entered as a continuous variable in weeks (0, 3, and 9). The mixed model employed the more conservative restricted maximum likelihood estimation and unstructured covariance matrix.

Results

Recruitment

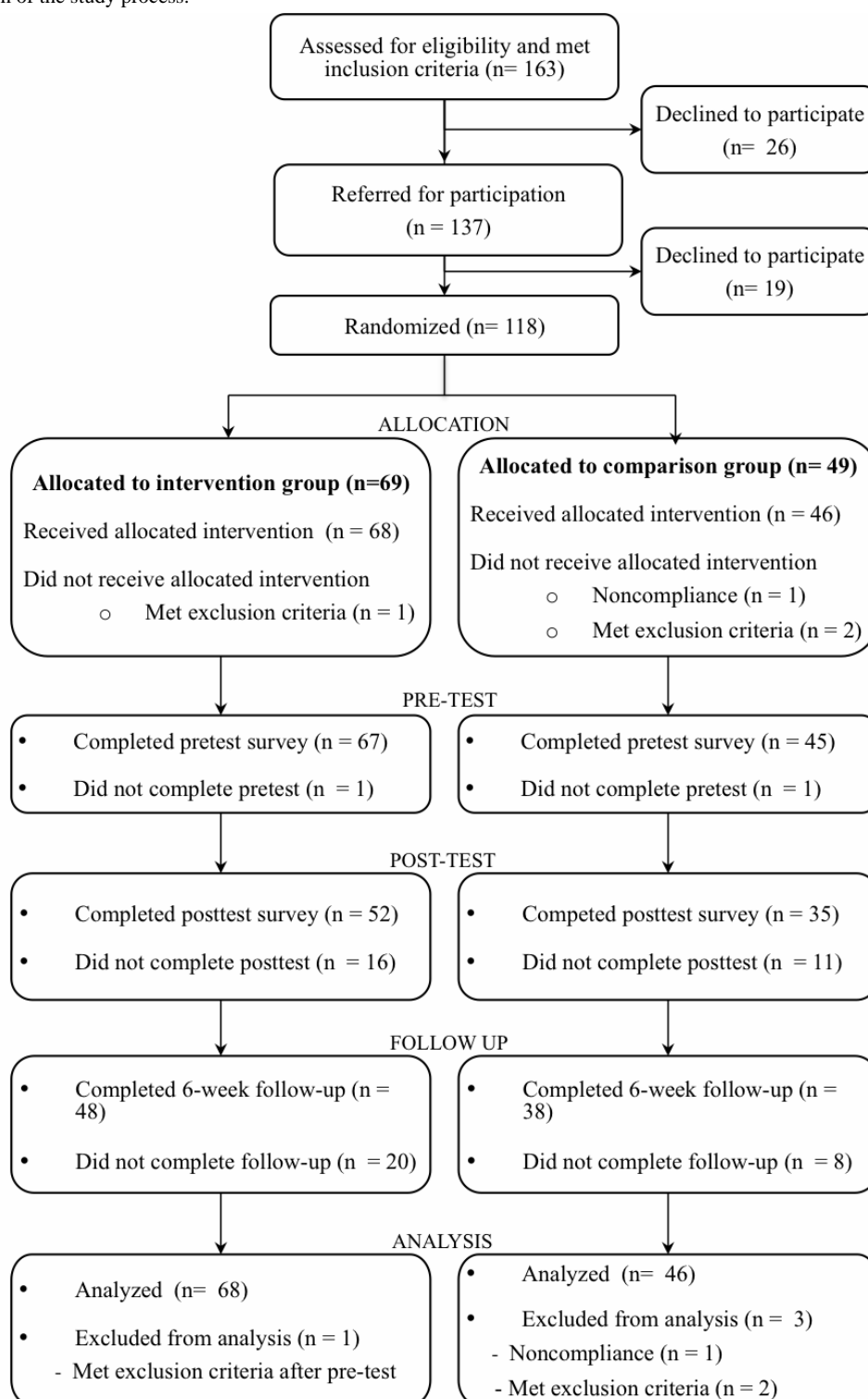
We collected data collection between April 16, 2009 and January 28, 2011. As seen in [Figure 1](#), 137 young people accepted the invitation to join the study, of whom 118 began the recruitment process. We excluded 4 participants after randomization (2 became too unwell to participate, 1 was incarcerated, and 1 gave invalid responses to all pretest measures), resulting in a final sample of 114 young people, which was sufficient to detect the primary aim of a medium-sized indirect effect [56]. Due to a failure to recruit the expected sample of 200 participants, there is a different number of participants in the comparison and intervention groups; however, a test of the binomial distribution indicated that this difference was not significant with 69 of 118 participants randomly allocated to the intervention group ($P = .08$). The total number of participants assessed for eligibility was difficult to establish, as the GPs rarely recorded information

about patients who met the inclusion criteria and were either not approached to participate or declined when invited to participate. Therefore, the number of patients assessed for eligibility presented in Figure 1 is likely to be underrepresented.

In total, 67% of participants (76/114) completed all questionnaires and 85% of participants (97/114) completed questionnaires at two or more time points. We conducted *t* tests and chi-square tests and found no significant differences

between participants who completed all questionnaires and those who missed questionnaires. Therefore, we included all 114 participants (68 in the intervention group and 46 in the comparison group) in analyses using the routines for missing data in the maximum likelihood estimation. Of 114 participants, 2 (2%) did not complete the pretest questionnaire but went on to complete the mobiletype entries and posttests. We could not contact 27 participants (24%) for the posttest questionnaire but they went on to complete the 6-week follow-up questionnaire.

Figure 1. Flow diagram of the study process.



Demographics

We found no statistically significant differences in demographic characteristics between the intervention and comparison groups on any pretest measures listed in Table 2, except that the intervention group reported significantly higher stress than the comparison group with a mean difference of 3.4 ($t_{109} = 2.06$, $P = .04$).

Participants in the intervention group completed an average of 3.3 (SD 1.42, range 1–8) mobiletype entries each day and completed the program from 1 to 34 days with a mean of 17.7

(SD 6.69) days completed. In the comparison group, participants completed an average of 4 (SD 1.77, range 1–12) mobiletype entries per day, and completed the program for 8 to 25 days with a mean of 16.8 (SD 4.03) days completed. The minimum dose of the program was completion of at least two mobiletype entries a day for at least 14 days. As Table 2 shows, 64 (56%) participants in intervention group and 36 (52.9%) in comparison received a minimum dose. Means, standard deviations and 95% confidence intervals are presented in Table 3 using the observed scores.

Table 2. General demographics of participants in the comparison and intervention groups (n = 114).

Characteristic	Comparison group	Intervention group	P value
Total number ^a , n (%)	49 (41.5%)	69 (58.5.0%)	.08
14 days completed ^b , n (%)	28 (59.6%)	36 (52.9%)	.33
Area , n (%)			
Melbourne	14 (30.4%)	28 (42.6%)	.27
Goulburn Valley	21 (45.7%)	21 (29.4%)	
Albury-Wodonga	11 (23.9%)	19 (27.9%)	
Male participants, n (%)	17 (37.0%)	15 (22.1%)	.08
Age (years), mean (SD)	17.4 (3.2)	18.5 (3.2)	.06
Ethnic identification ^c , n (%)	4 (9.1%)	10 (22.7%)	.37
Employment , n (%)			
Employed	7 (15.2%)	18 (26.5%)	.21
Unemployed	4 (8.7%)	9 (13.2%)	
Student	35 (76.1%)	41 (60.3%)	
Drug-related items ^c			
Ever had alcohol	38 (86.4%)	59 (88.1%)	.79
Ever been drunk	31 (70.5%)	53 (79.1%)	.30
Ever had a cigarette	25 (56.8%)	38 (56.7%)	.99
Ever tried marijuana	18 (40.9%)	33 (49.3%)	.39
Ever tried other ^d drugs	10 (22.7%)	26 (38.8%)	.08
Pretest DASS ^{c,e} , mean (SD)			
Depression	19.4 (10.8)	20.4 (11.0)	.63
Anxiety	11.0 (8.0)	14.1 (9.7)	.08
Stress	16.9 (7.9)	20.3 (8.9)	.04

^a Binomial test on number at randomization (n = 118).

^b Completed mobiletype entries at least twice daily for 14 days.

^c Observed means (n = 111).

^d Sedatives, tranquilizers, amphetamines, analgesics, inhalants, cocaine, LSD, and heroin.

^e Depression Anxiety Stress Scale.

Table 3. Descriptive statistics for the intervention and comparison groups' scores^a on depression and emotional self-awareness at pretest, posttest, and 6-week follow-up.

	Comparison group			Intervention group		
	n ^b	Mean (SD)	95% CI	n ^b	Mean (SD)	95% CI
Depression						
Pretest	44	19.4 (10.9)	16.1–22.7	67	20.4 (11.0)	17.8–23.1
Posttest	33	15.2 (8.9)	12.1–18.3	50	16.3 (10.8)	13.3–19.4
6-week follow-up	36	12.5 (11.8)	8.5–16.5	50	13.5 (10.5)	10.5–16.5
Emotional self-awareness						
Pretest	44	61.1 (11.9)	57.4–64.7	67	61.6 (12.1)	58.7–64.6
Posttest	32	63.1 (11.1)	59.1–67.1	46	64.7 (10.9)	60.9–67.4
6-week follow-up	35	62.2 (11.6)	58.2–66.1	47	68.9 (11.2)	65.5–72.1
Rumination						
Pretest	44	12.8 (3.16)	11.9–13.8	67	14.0 (3.43)	13.2–14.9
Posttest	33	12.2 (3.57)	10.9–13.4	46	12.4 (3.57)	11.3–13.4
6-week follow-up	35	11.2 (3.67)	10.0–12.5	48	11.7 (3.62)	10.7–12.8

^a Observed scores.^b Number of participants used to calculate the mean, standard deviation (SD), and 95% confidence interval (CI).

Parallel Process Latent Growth Curve Model

The path diagram in [Figure 2](#) shows the parallel process LGCM used to test the indirect pathway of the mobiletype program used on depressive symptoms via ESA. The fit indices for [Figure 2](#) suggest that the model is a good fit for the data ($\chi^2_6 = 11.3$, $P = .08$, comparative fit index = .958, Tucker-Lewis Index = .854, root mean square error of approximation = .088, $P = .18$, standardized root mean square residual = .040).

All possible pathways were calculated between group and the four latent variables as illustrated in [Figure 2](#). The first step to testing mediation was conducted by regressing the slope of ESA onto group. The positive coefficient indicates that there was a greater increase in ESA across time in the intervention group than in the comparison group. This coefficient was statistically significant, as the 95% bias-corrected bootstrap CIs of group on ESA did not contain zero (see [Table 4](#)). The second step to testing mediation was conducted by regressing the slope of ESA onto the slope of depression. The negative coefficient here indicates that an increase in ESA led to a decrease in depressive symptoms. This pathway was also statistically significant as

indicated by the 95% bias-corrected bootstrap CIs not containing zero (see [Table 4](#)). All other pathways illustrated in [Figure 2](#) were tested and were not statistically significant.

The indirect effect of group on the slope of depression via the slope of ESA reported in [Table 4](#) is statistically significant (95% bootstrap CI did not contain zero) with an unstandardized estimate of -0.608 , indicating that participants in the intervention group had a decrease in depressive symptoms via the slope of ESA when compared with the comparison group. There was no significant direct effect from group to depressive symptoms based on the 95% bootstrap CI, indicating that the intervention did not directly decrease depressive symptoms.

[Figure 3](#) presents the relationship between the change in depressive symptoms and the change in ESA for both the intervention and comparison groups.

There was a negative relationship between changes in ESA and changes in depressive symptoms for both groups; accordingly, an increase in ESA was associated with a decrease in depressive symptoms as seen in [Figure 3](#). The intervention group had larger increases in ESA than did the comparison group, which was associated with a greater decrease in depressive symptoms.

Table 4. Coefficients and bias-corrected bootstrapping confidence intervals (CIs) of the parallel process latent growth curve model.

Effect	Point estimate	95% Bootstrap CI
Direct effect on the slope of depression		
Slope of ESA	-0.902 ^a	-6.209 to -0.052
Intercept of depression	-0.052	-0.126 to 0.012
Intercept of ESA	-0.026	-0.315 to 0.027
Group	0.587	-0.114 to 5.072
Direct effect on the slope of ESA		
Intercept of ESA	-0.044	-0.083 to 0.162
Intercept of depression	-0.003	-0.038 to 0.063
Group	0.676 ^a	0.019 to 1.231
Direct effect of group on intercept		
ESA	0.439	-3.904 to 4.562
Depression	1.018	-2.980 to 5.208
Indirect effect of group on slope of depression		
Via the slope of ESA	-0.610 ^a	-5.596 to -0.003
Via the intercept of ESA	-0.012	-0.526 to 0.105
Total indirect effect	-0.621 ^b	-6.269 to -0.036

^a Confidence interval does not contain zero.

Figure 2. Parallel process latent growth curve model of the group effect on the growth of depressive symptoms via the growth of emotional self-awareness (ESA). Unstandardized estimates reported; boldface lines represent statistically significant pathways. ^aTime interval from pretest by week. $DEP_{pre, post, 6-week}$, $ESA_{pre, post, 6-week}$ = the observed score of the Depression Anxiety Stress Scale depression subscale and ESA scale at pretest, posttest, and 6-week follow-up, respectively; Group = intervention program condition; i_{DEP} = latent intercept of depression; i_{ESA} = latent intercept of ESA; s_{DEP} = latent slope of depression; s_{ESA} = latent slope of ESA.

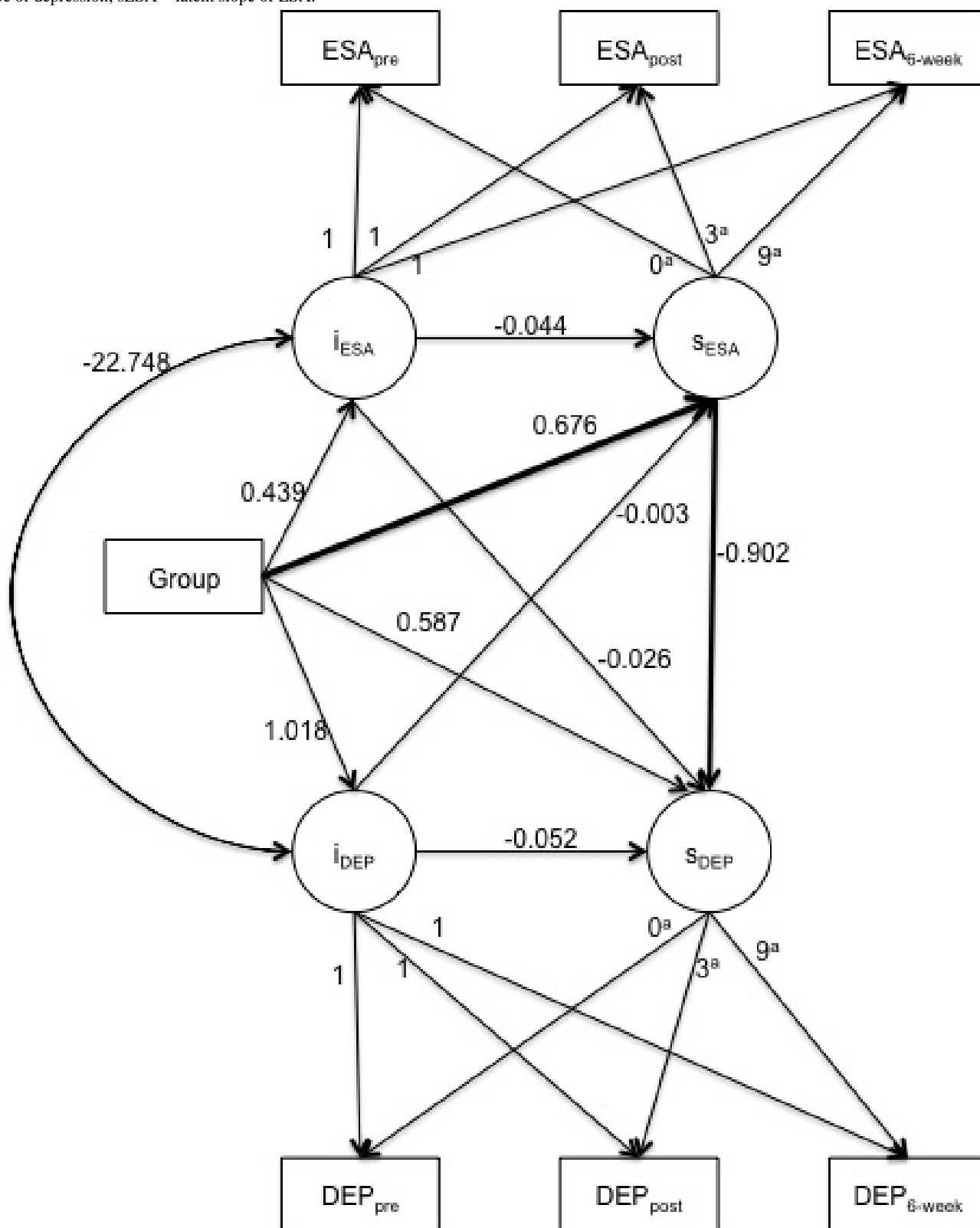
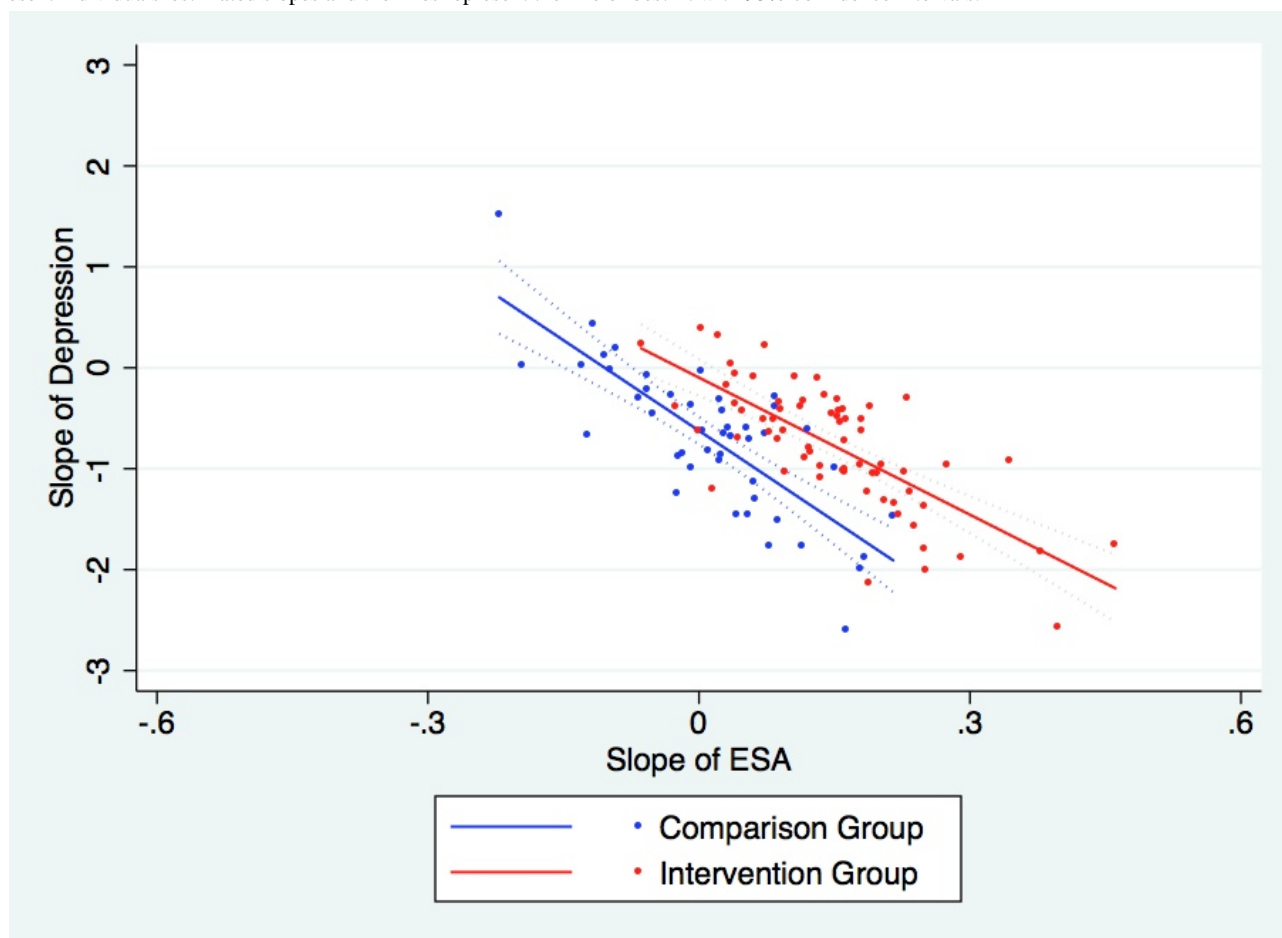


Figure 3. The relationship between self-monitoring, the slope of depressive symptoms, and the slope of emotional self-awareness (ESA). The points represent individuals' estimated slopes and the lines represent the line of best fit with 95% confidence intervals.



Indirect Effect Sizes

The unstandardized indirect effect size can be interpreted on the DASS depression subscale (on a scale from 0 to 42), indicating that the intervention group is estimated to have a linear decrease in depressive symptoms of .688 per week (95% CI $-.962$ to $-.487$) indirectly through mediation of the linear slope of ESA when compared with the comparison group. The proportion of the maximum possible indirect effect has similar properties to, and can be interpreted on a similar scale to, Cohen's r^2 [67]; therefore, the estimated κ^2 of .54 (95% CI .426–.640) indicates a large size of indirect effect.

Rumination

We conducted a secondary analysis to determine whether rumination decreased over time between groups. The mixed model analysis of the Ruminative Response Scale brooding subscale showed a significant main effect of time ($\beta = -0.16$, $P = .02$), indicating an overall decrease in rumination over time for both groups. There was no significant main difference between groups ($\beta = 1.01$, $P = .11$), nor was the interaction between group and time significant ($\beta = -0.07$, $P = .39$).

Discussion

The current study examined the use of a mobile phone self-monitoring program on ESA with young people who had mild or more depressive symptoms, and supported the

hypothesis that self-monitoring mood, stress, and coping strategies increases awareness of emotions. The second hypothesis that an increase in ESA would predict a decrease in depressive symptoms was also supported. Based on Preacher and Kelley's proportion of the maximum possible indirect effect [67], there was a large effect of the intervention program on depressive symptoms indirectly via ESA.

This study supports previous research suggesting that simple self-monitoring techniques effectively increase self-awareness, in this case, awareness of one's own emotions [34,68]. Metacognitions, such as self-awareness, are developed during early adolescence [35], and interventions can be developed that target young people's ability to recognize emotions, identify emotional states, understand the contextualization of emotions, communicate this emotional knowledge, and plan and make constructive decisions about emotions. Increasing ESA is a core process in the early stages of therapy [32]. The current study demonstrates the potential for targeting ESA in first-step intervention strategies for young people with mild or more depressive symptoms. This randomized controlled trial was conducted with a view of representing a wide variety of young people who visit GPs with a range of medical and psychological problems and severity of problems. Therefore, the results of this study are applicable to this age group in general.

Self-monitoring techniques may provide an alternative to watchful waiting as a first-step intervention in the stepped-care

approach. Mobile phones are ideally suited to this purpose, as the mobiletype program can be downloaded to patients' own mobile phones to help young people understand and manage mild depressive symptoms. Detailed information about patients' mental health in recent weeks is then uploaded to GPs in an easy-to-read format, saving time in appointments and allowing progression to more intensive second-step interventions when needed. Young people often do not recognize mental health problems [69] and instead attend GP clinics for somatic complaints rather than mental health symptoms [70]. Using self-monitoring techniques with young people who present with underlying somatic complaints may increase their ESA and help young people initiate treatment for depression.

Our secondary hypothesis that the intervention program participants would have a decrease in rumination when compared with those in the comparison program was not supported. Further research is needed to determine whether there is an inverse relationship between rumination and ESA. Nevertheless, rumination decreased over time as did depressive symptoms [47], further supporting the positive relationship between depression and rumination.

Primary care is a particularly difficult setting in which to conduct randomized controlled studies [71,72]. In this study, recruitment was delayed and ceased before the anticipated sample size was recruited due to the H1N1 influenza pandemic, first detected in Melbourne on May 9, 2009 [57], and during school holidays. These delays resulted in uneven numbers in the two groups. We strongly recommend allocating extra time for recruitment in primary care compared with other settings, particularly in youth-focused studies. The intervention program had no direct effect on depressive symptoms. One interpretation of these results is that there was reduced power for a direct effect given that depressive symptoms decreased significantly over time in both groups. It is possible that both groups had a decrease in depressive symptoms due to the resources, training, and support given to the GPs; however, a larger sample size, or a wait-list control group, would be needed to determine whether

depressive symptoms differed between the groups [56,62]. Finally, Reid et al [47] detail other limitations: a cluster randomized controlled trial, in which GPs rather than patients are randomly allocated, may have been more appropriate but was rejected during the study design due to the difficulty in blinding GPs and participants to the randomization procedure; and participant heterogeneity in illness type, severity, and familiarity with their GP due to broad inclusion criteria needed in an effectiveness trial is likely to have reduced the overall power of the study.

To our knowledge, this is the first randomized controlled trial examining the use of a mobile phone self-monitoring program as an intervention tool for young people with depressive symptoms. Self-monitoring was shown to effectively decrease depression via the mechanism of ESA, suggesting that self-monitoring programs that focus on increasing ESA may provide a useful framework for first-step care in depression. The program provided GPs with information about a young person's daily activities and can be used to detect early signs of mental health problems, such as elevated negative mood, stress and causes of stress, maladaptive coping strategies, isolation from peers, diet, and exercise, as well as other risk and protective factors. The mobile phone self-monitoring program has the advantage of being low cost, quick, and easy to use.

In summary, mobile phones are well suited to first-step interventions, providing an alternative to watchful waiting and allowing young people to provide accurate information to their GPs about their mood and stress [21], as well as shortening the length of time it would take to relay this information to GPs in a usual appointment. Mobile phone self-monitoring programs should be considered as a first-step low-cost intervention with young people who are at risk of mental health problems. Self-monitoring has the advantages of helping young people increase their ESA while gaining more information about their mental health symptoms in order to direct them to the best intervention.

Acknowledgments

This study was supported by research grants from the Telstra Foundation and The Shepherd Foundation, and the Victorian Government's Operational Infrastructure Support Program supported infrastructure. Sylvia Kauer received an Ian Scott PhD scholarship funded by Australian Rotary Health. Be.Interactive provided telecommunication support. Telstra Corporation provided prepaid SIM cards and telecommunication support. Sony Ericsson provided 40 mobile phones for use in this study. The Statistical Mediation and Moderation Analysis discussion forum on Facebook, led by Kristopher J Preacher and Andrew F Hayes, provided support for the mediation analyses conducted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Emotional self-awareness scale.

[PDF File (Adobe PDF File), 26KB - [jmir_v14i3e67_app1.pdf](#)]

Multimedia Appendix 2

Indirect effect size estimates of the the parallel process latent growth curve model for group on depressive symptoms via the pathway of emotional self-awareness.

[[PDF File \(Adobe PDF File\), 42KB - jmir_v14i3e67_app2.pdf](#)]

References

1. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997 May 24;349(9064):1498-1504. [doi: [10.1016/S0140-6736\(96\)07492-2](#)] [Medline: [9167458](#)]
2. Kessler RC, Avenevoli S, Ries Merikangas K. Mood disorders in children and adolescents: an epidemiologic perspective. *Biol Psychiatry* 2001 Jun 15;49(12):1002-1014. [Medline: [11430842](#)]
3. Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994 Jul;151(7):979-986. [Medline: [8010383](#)]
4. Andrews G, Szabo M, Burns J. Preventing major depression in young people. *Br J Psychiatry* 2002 Dec;181:460-462 [[FREE Full text](#)] [Medline: [12456512](#)]
5. Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *J Abnorm Psychol* 1993 Feb;102(1):133-144. [Medline: [8436689](#)]
6. Rushton JL, Forcier M, Schectman RM. Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. *J Am Acad Child Adolesc Psychiatry* 2002 Feb;41(2):199-205. [doi: [10.1097/00004583-200202000-00014](#)] [Medline: [11837410](#)]
7. Jane Costello E, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? *J Child Psychol Psychiatry* 2006 Dec;47(12):1263-1271. [doi: [10.1111/j.1469-7610.2006.01682.x](#)] [Medline: [17176381](#)]
8. Andrews G, Sanderson K, Corry J, Lapsley HM. Using epidemiological data to model efficiency in reducing the burden of depression. *J Ment Health Policy Econ* 2000 Dec 1;3(4):175-186. [Medline: [11967454](#)]
9. Jorm AF, Griffiths KM. Population promotion of informal self-help strategies for early intervention against depression and anxiety. *Psychol Med* 2006 Jan;36(1):3-6. [doi: [10.1017/S0033291705005659](#)] [Medline: [16356291](#)]
10. Geisner IM. Alternative brief interventions for mild depression. *Psychiatr Times* 2006;23(11).
11. van Straten A, Seekles W, van 't Veer-Tazelaar NJ, Beekman AT, Cuijpers P. Stepped care for depression in primary care: what should be offered and how? *Med J Aust* 2010 Jun 7;192(11 Suppl):S36-S39. [Medline: [20528706](#)]
12. Lynch FL, Hornbrook M, Clarke GN, Perrin N, Polen MR, O'Connor E, et al. Cost-effectiveness of an intervention to prevent depression in at-risk teens. *Arch Gen Psychiatry* 2005 Nov;62(11):1241-1248 [[FREE Full text](#)] [doi: [10.1001/archpsyc.62.11.1241](#)] [Medline: [16275811](#)]
13. Spence S, Burns J, Boucher S, Glover S, Graetz B, Kay D, et al. The beyondblue Schools Research Initiative: conceptual framework and intervention. *Australas Psychiatry* 2005 Jun;13(2):159-164. [doi: [10.1111/j.1440-1665.2005.02180.x](#)] [Medline: [15948913](#)]
14. Spence SH, Sheffield JK, Donovan CL. Preventing adolescent depression: an evaluation of the problem solving for life program. *J Consult Clin Psychol* 2003 Feb;71(1):3-13. [Medline: [12602420](#)]
15. Britt H, Valenti L, Miller G. Time for care. Length of general practice consultations in Australia. *Aust Fam Physician* 2002 Sep;31(9):876-880. [Medline: [12402711](#)]
16. Haller DM, Sanci LA, Sawyer SM, Patton GC. The identification of young people's emotional distress: a study in primary care. *Br J Gen Pract* 2009 Mar;59(560):e61-e70. [doi: [10.3399/bjgp09X419510](#)] [Medline: [19275825](#)]
17. Stice E, Shaw H, Bohon C, Marti CN, Rohde P. A meta-analytic review of depression prevention programs for children and adolescents: factors that predict magnitude of intervention effects. *J Consult Clin Psychol* 2009 Jun;77(3):486-503. [doi: [10.1037/a0015168](#)] [Medline: [19485590](#)]
18. Walsh SP, White KM, Young RM. Over-connected? A qualitative exploration of the relationship between Australian youth and their mobile phones. *J Adolesc* 2008 Feb;31(1):77-92. [doi: [10.1016/j.adolescence.2007.04.004](#)] [Medline: [17560644](#)]
19. Zhao S, Grasmuck S, Martin J. Identity construction on Facebook: digital empowerment in anchored relationships. *Comput Hum Behav* 2008;24:1816-1836. [doi: [10.1016/j.chb.2008.02.012](#)]
20. Hollon SD, Muñoz RF, Barlow DH, Beardslee WR, Bell CC, Bernal G, et al. Psychosocial intervention development for the prevention and treatment of depression: promoting innovation and increasing access. *Biol Psychiatry* 2002 Sep 15;52(6):610-630. [Medline: [12361671](#)]
21. Reid SC, Kauer SD, Dudgeon P, Sanci LA, Shrier LA, Patton GC. A mobile phone program to track young people's experiences of mood, stress and coping. Development and testing of the mobiletype program. *Soc Psychiatry Psychiatr Epidemiol* 2009 Jun;44(6):501-507. [doi: [10.1007/s00127-008-0455-5](#)] [Medline: [19011723](#)]
22. Silk JS, Steinberg L, Morris AS. Adolescents' emotion regulation in daily life: links to depressive symptoms and problem behavior. *Child Dev* 2003;74(6):1869-1880. [Medline: [14669901](#)]

23. Kauer SD, Reid SC, Sanci L, Patton GC. Investigating the utility of mobile phones for collecting data about adolescent alcohol use and related mood, stress and coping behaviours: lessons and recommendations. *Drug Alcohol Rev* 2009 Jan;28(1):25-30. [doi: [10.1111/j.1465-3362.2008.00002.x](https://doi.org/10.1111/j.1465-3362.2008.00002.x)] [Medline: [19320672](https://pubmed.ncbi.nlm.nih.gov/19320672/)]
24. Stone AA, Shiffman S, Atienza AA, Nebeling L. Historical roots and rationale of ecological momentary assessment (EMA). In: Stone AA, Shiffman S, Atienza AA, Nebeling L, editors. *The Science of Real-Time Data Capture: Self-reports in Health Research*. New York, NY: Oxford University Press; 2007:3-10.
25. Kauer S, Reid S, Jackson H, Jorm T. The phenomenon of insight and preliminary findings for preventing depression: self-monitoring via mobile phones. *Aust N Z J Psychiatry* 2007;41(s2):A469. [doi: [10.1080/00048670701757092](https://doi.org/10.1080/00048670701757092)]
26. Kazantzis N. Power to detect homework effects in psychotherapy outcome research. *J Consult Clin Psychol* 2000 Feb;68(1):166-170. [Medline: [10710851](https://pubmed.ncbi.nlm.nih.gov/10710851/)]
27. Abueg FR, Colletti G, Kopel SA. A study of reactivity: the effects of increased relevance and saliency of self-monitored smoking through enhanced carbon monoxide feedback. *Cogn Ther Res* 1985;9:321-333. [doi: [10.1007/BF01183851](https://doi.org/10.1007/BF01183851)]
28. Ewart CK. Self-observation in natural environments: reactive effects of behavior desirability and goal-setting. *Cogn Ther Res* 1978;2:39-56. [doi: [10.1007/BF01172511](https://doi.org/10.1007/BF01172511)]
29. Kazdin AE. Reactive self-monitoring: the effects of response desirability, goal setting, and feedback. *J Consult Clin Psychol* 1974 Oct;42(5):704-716. [Medline: [4427011](https://pubmed.ncbi.nlm.nih.gov/4427011/)]
30. Cheong J, Mackinnon DP, Khoo ST. Investigation of mediational processes using parallel process latent growth curve modeling. *Struct Equ Modeling* 2003 Apr 1;10(2):238-262. [doi: [10.1207/S15328007SEM1002_5](https://doi.org/10.1207/S15328007SEM1002_5)] [Medline: [20157639](https://pubmed.ncbi.nlm.nih.gov/20157639/)]
31. Sheffield JK, Spence SH, Rapee RM, Kowalenko N, Wignall A, Davis A, et al. Evaluation of universal, indicated, and combined cognitive-behavioral approaches to the prevention of depression among adolescents. *J Consult Clin Psychol* 2006 Feb;74(1):66-79. [doi: [10.1037/0022-006X.74.1.66](https://doi.org/10.1037/0022-006X.74.1.66)] [Medline: [16551144](https://pubmed.ncbi.nlm.nih.gov/16551144/)]
32. Prochaska JO, DiClemente CC. Transtheoretical therapy: towards a more integrative model of change. *Psychotherapy (Chic)* 1982;19(3):276-288. [doi: [10.1037/h0088437](https://doi.org/10.1037/h0088437)]
33. Beck AT. The past and future of cognitive therapy. *J Psychother Pract Res* 1997;6(4):276-284 [FREE Full text] [Medline: [9292441](https://pubmed.ncbi.nlm.nih.gov/9292441/)]
34. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res* 2010;12(2):e10 [FREE Full text] [doi: [10.2196/jmir.1371](https://doi.org/10.2196/jmir.1371)] [Medline: [20439251](https://pubmed.ncbi.nlm.nih.gov/20439251/)]
35. Saari C. Emotional competence and self-regulation in childhood. In: Salovey P, Sluyter DJ, editors. *Emotional Development and Emotional Intelligence: Educational Implications*. New York, NY: Basic Books; 1997:35-66.
36. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. *J Pers Soc Psychol* 2003 Apr;84(4):822-848. [Medline: [12703651](https://pubmed.ncbi.nlm.nih.gov/12703651/)]
37. Chambers R, Lo BCY, Allen NB. The impact of intensive mindfulness training on attentional control, cognitive style, and affect. *Cognit Ther Res* 2008;32:303-322. [doi: [10.1007/s10608-007-9119-0](https://doi.org/10.1007/s10608-007-9119-0)]
38. Beitman BD, Soth AM. Activation of self-observation: a core process among the psychotherapies. *J Psychother Integr* 2006;16:383-397. [doi: [10.1037/1053-0479.16.4.383](https://doi.org/10.1037/1053-0479.16.4.383)]
39. Lane RD, Schwartz GE. Levels of emotional awareness: a cognitive-developmental theory and its application to psychopathology. *Am J Psychiatry* 1987 Feb;144(2):133-143. [Medline: [3812780](https://pubmed.ncbi.nlm.nih.gov/3812780/)]
40. Swinkels A, Giuliano TA. The measurement and conceptualisation of mood awareness: monitoring and labeling one's mood states. *Pers Soc Psychol Bull* 1995;21:934-949. [doi: [10.1177/0146167295219008](https://doi.org/10.1177/0146167295219008)]
41. Watkins ER. Constructive and unconstructive repetitive thought. *Psychol Bull* 2008 Mar;134(2):163-206. [doi: [10.1037/0033-2909.134.2.163](https://doi.org/10.1037/0033-2909.134.2.163)] [Medline: [18298268](https://pubmed.ncbi.nlm.nih.gov/18298268/)]
42. Epstein S. Integration of the cognitive and the psychodynamic unconscious. *Am Psychol* 1994 Aug;49(8):709-724. [Medline: [8092614](https://pubmed.ncbi.nlm.nih.gov/8092614/)]
43. Lyke JA. Insight, but not self-reflection, is related to subjective well-being. *Pers Individ Diff* 2009;46:66-70. [doi: [10.1016/j.paid.2008.09.010](https://doi.org/10.1016/j.paid.2008.09.010)]
44. Segerstrom SC, Stanton AL, Alden LE, Shortridge BE. A multidimensional structure for repetitive thought: what's on your mind, and how, and how much? *J Pers Soc Psychol* 2003 Nov;85(5):909-921. [doi: [10.1037/0022-3514.85.5.909](https://doi.org/10.1037/0022-3514.85.5.909)] [Medline: [14599253](https://pubmed.ncbi.nlm.nih.gov/14599253/)]
45. Grosse Holtforth M, Casonguay LG, Boswell JF, Wilson LA, Kakouros AA, Borkovec TD. Insight in cognitive-behavioral therapy. In: Castonguay LG, Hill CE, editors. *Insight in Psychotherapy*. Washington, DC: American Psychological Association; 2007:57-80.
46. Mayer JD, Stevens AA. An emerging understanding of the reflective (meta-) experience of mood. *J Res Pers* 1994;28:351-373. [doi: [10.1006/jrpe.1994.1025](https://doi.org/10.1006/jrpe.1994.1025)]
47. Reid SC, Kauer SD, Hearps SJ, Crooke AH, Khor AS, Sanci LA, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: a randomised controlled trial. *BMC Fam Pract* 2011;12:131 [FREE Full text] [doi: [10.1186/1471-2296-12-131](https://doi.org/10.1186/1471-2296-12-131)] [Medline: [22123031](https://pubmed.ncbi.nlm.nih.gov/22123031/)]
48. von Soest T, Hagtvet K. Mediation analysis in a latent growth curve modeling framework. *Struct Equ Modeling* 2011;18(2):289-314. [doi: [10.1080/10705511.2011.557344](https://doi.org/10.1080/10705511.2011.557344)]

49. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;8:18-27 [[FREE Full text](#)] [doi: [10.1186/1741-7015-8-18](https://doi.org/10.1186/1741-7015-8-18)] [Medline: [20334633](#)]
50. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). *Aust N Z J Public Health* 2001 Dec;25(6):494-497. [Medline: [11824981](#)]
51. Lindquist R, Wyman JF, Talley KM, Findorff MJ, Gross CR. Design of control-group conditions in clinical trials of behavioral interventions. *J Nurs Scholarsh* 2007;39(3):214-221. [doi: [10.1111/j.1547-5069.2007.00171.x](https://doi.org/10.1111/j.1547-5069.2007.00171.x)] [Medline: [17760793](#)]
52. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther* 1995 Mar;33(3):335-343. [Medline: [7726811](#)]
53. Grant AM, Franklin J, Langford P. The Self-Reflection and Insight Scale: a new measure of private self-consciousness. *Soc Behav Pers* 2002;30(8):821-836. [doi: [10.2224/sbp.2002.30.8.821](https://doi.org/10.2224/sbp.2002.30.8.821)]
54. Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: a psychometric analysis. *Cogn Ther Res* 2003;27:247-259. [doi: [10.1023/A:1023910315561](https://doi.org/10.1023/A:1023910315561)]
55. Cohen J. A power primer. *Psychol Bull* 1992 Jul;112(1):155-159. [Medline: [19565683](#)]
56. Fritz MS, Mackinnon DP. Required sample size to detect the mediated effect. *Psychol Sci* 2007 Mar;18(3):233-239. [doi: [10.1111/j.1467-9280.2007.01882.x](https://doi.org/10.1111/j.1467-9280.2007.01882.x)] [Medline: [17444920](#)]
57. Denholm JT, Gordon CL, Johnson PD, Hewagama SS, Stuart RL, Aboltins C, et al. Hospitalised adult patients with pandemic (H1N1) 2009 influenza in Melbourne, Australia. *Med J Aust* 2010 Jan 18;192(2):84-86. [Medline: [20078408](#)]
58. Hedeker D, Gibbons RD. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychol Methods* 1997;2(1):64-78.
59. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986 Dec;51(6):1173-1182. [Medline: [3806354](#)]
60. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007;58:593-614. [doi: [10.1146/annurev.psych.58.110405.085542](https://doi.org/10.1146/annurev.psych.58.110405.085542)] [Medline: [16968208](#)]
61. Selig JP, Preacher KJ. Mediation models for longitudinal data in developmental research. *Res Hum Dev* 2009;6(2-3):144-164. [doi: [10.1080/15427600902911247](https://doi.org/10.1080/15427600902911247)]
62. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr* 2009;76(4):408-420. [doi: [10.1080/03637750903310360](https://doi.org/10.1080/03637750903310360)]
63. Jackson DL, Gillaspay JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol Methods* 2009 Mar;14(1):6-23. [doi: [10.1037/a0014694](https://doi.org/10.1037/a0014694)] [Medline: [19271845](#)]
64. Schermelleh-Engel K, Moosbrugger H, Muller H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res Online* 2003;8(2):23-74.
65. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods* 2008 Aug;40(3):879-891. [Medline: [18697684](#)]
66. Taylor AB, MacKinnon DP, Tein JY. Test of the three-path mediated effect. *Organ Res Methods* 2008;11(2):241-269. [doi: [10.1177/1094428107300344](https://doi.org/10.1177/1094428107300344)]
67. Preacher KJ, Kelley K. Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychol Methods* 2011 Jun;16(2):93-115. [doi: [10.1037/a0022658](https://doi.org/10.1037/a0022658)] [Medline: [21500915](#)]
68. Pennebaker JW, Zech E, Rime B. Disclosing and sharing emotion: psychological, social and health consequences. In: Stroebe MS, Stroebe W, Hansson RO, Schut H, editors. *Handbook of Bereavement Research: Consequences, Coping and Care*. Washington, DC: American Psychological Association; 2001:517-539.
69. Rickwood D, Cavanagh S, Curtis L, Sakrouge R. Educating young people about mental health and mental illness: evaluating a school-based programme. *Int J Ment Health Promot* 2004;6(4):23-32.
70. Tylee A, Haller DM, Graham T, Churchill R, Sanci LA. Youth-friendly primary-care services: how are we doing and what more needs to be done? *Lancet* 2007 May 5;369(9572):1565-1573. [doi: [10.1016/S0140-6736\(07\)60371-7](https://doi.org/10.1016/S0140-6736(07)60371-7)] [Medline: [17482988](#)]
71. Page MJ, French SD, McKenzie JE, O'Connor DA, Green SE. Recruitment difficulties in a primary care cluster randomised trial: investigating factors contributing to general practitioners' recruitment of patients. *BMC Med Res Methodol* 2011;11:35 [[FREE Full text](#)] [doi: [10.1186/1471-2288-11-35](https://doi.org/10.1186/1471-2288-11-35)] [Medline: [21453543](#)]
72. Williamson MK, Pirkis J, Pfaff JJ, Tyson O, Sim M, Kerse N, et al. Recruiting and retaining GPs and patients in intervention studies: the DEPS-GP project as a case study. *BMC Med Res Methodol* 2007;7:42 [[FREE Full text](#)] [doi: [10.1186/1471-2288-7-42](https://doi.org/10.1186/1471-2288-7-42)] [Medline: [17875219](#)]

Abbreviations

DASS: Depression Anxiety Stress Scale
ESA: emotional self-awareness
GP: general practitioner
LGCM: latent growth curve model

Edited by D Mohr, A Caele, H Christensen; submitted 08.06.11; peer-reviewed by G Andrews, G Clarke; comments to author 06.07.11; revised version received 18.08.11; accepted 23.08.11; published 25.06.12.

Please cite as:

Kauer SD, Reid SC, Crooke AHD, Khor A, Hearps SJC, Jorm AF, Sancí L, Patton G

Self-monitoring Using Mobile Phones in the Early Stages of Adolescent Depression: Randomized Controlled Trial

J Med Internet Res 2012;14(3):e67

URL: <http://www.jmir.org/2012/3/e67/>

doi: [10.2196/jmir.1858](https://doi.org/10.2196/jmir.1858)

PMID: [22732135](https://pubmed.ncbi.nlm.nih.gov/22732135/)

©Sylvia Deidre Kauer, Sophie Caroline Reid, Alexander Hew Dale Crooke, Angela Khor, Stephen John Charles Hearps, Anthony Francis Jorm, Lena Sancí, George Patton. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 25.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Individually Tailored Internet-Based Treatment for Young Adults and Adults With Panic Attacks: Randomized Controlled Trial

Kristin Silfvernagel¹, MSc; Per Carlbring², MSc, PhD; Julia Kabo¹, MSc; Sara Edström¹, MSc; Jenny Eriksson¹, MSc; Lisa Månson¹, MSc; Gerhard Andersson^{1,3,4}, BSc, MSc, PhD

¹Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

²Department of Psychology, Umeå University, Umeå, Sweden

³Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

⁴Department of Clinical Neuroscience, Psychiatry Section, Karolinska Institute, Stockholm, Sweden

Corresponding Author:

Kristin Silfvernagel, MSc

Department of Behavioural Sciences and Learning

Linköping University

IBL

Linköping, 58183

Sweden

Phone: 46 730342457

Fax: 46 13282145

Email: kristin.silfvernagel@liu.se

Abstract

Background: Previous studies on Internet-based treatment with minimal therapist guidance have shown promising results for several specific diagnoses.

Objective: To (1) investigate the effects of a tailored, therapist-guided, Internet-based treatment for individuals with reoccurring panic attacks, and (2) to examine whether people in different age groups (18–30 years and 31–45 years) would respond differently to the treatment.

Methods: We recruited 149 participants from an online list of individuals having expressed an interest in Internet treatment. Screening consisted of online questionnaires followed by a telephone interview. A total of 57 participants were included after a semistructured diagnostic interview, and they were randomly assigned to an 8-week treatment program ($n = 29$) or to a control condition ($n = 28$). Treatment consisted of individually prescribed cognitive behavior therapy text modules in conjunction with online therapist guidance. The control group consisted of people on a waitlist who later received treatment.

Results: All dependent measures improved significantly immediately following treatment and at the 12-month follow-up. The between-group effect size on the primary outcome measure, the Panic Disorder Severity Scale, was $d = 1.41$ (95% confidence interval 0.81–1.95) at posttreatment. The within-group effect size from pretreatment to 12-month follow-up was $d = 1.66$ (95% confidence interval 1.14–2.35). Age group had no effect, suggesting that age did not influence the outcome.

Conclusions: Tailoring an Internet-based treatment can be a feasible approach in the treatment of panic symptoms and comorbid anxiety and depressive symptoms. Younger adults benefit as much as adults over 30 years and up to 45 years of age.

Trial Registration: Clinicaltrials.gov NCT01296321; <http://www.clinicaltrials.gov/ct2/show/NCT01296321> (Archived by WebCite at <http://www.webcitation.org/65wddsqlL>)

(*J Med Internet Res* 2012;14(3):e65) doi:[10.2196/jmir.1853](https://doi.org/10.2196/jmir.1853)

KEYWORDS

Anxiety; depression; effectiveness; Internet-based treatment; cognitive behavior therapy

Introduction

Internet-based cognitive behavior therapy (iCBT) [1] has emerged as a novel evidence-based treatment for anxiety and mood disorders [2,3]. Guided iCBT has been found to be effective for numerous specific disorders such as panic disorder [4,5], generalized anxiety disorder [6,7], social anxiety disorder [8-10], posttraumatic stress disorder [11], specific phobia [12], major depression [3], stress [13], and somatic health problems such as irritable bowel syndrome [14]. One limitation of many previous iCBT trials on anxiety and depression is that they targeted specific disorders and that comorbid disorders could either be affected without being directly addressed or remain undetected and unchanged [15]. Another potential limitation of previous iCBT trials is that participants with comorbid anxiety and depression disorders often are excluded [16]. Structured or diagnosis-specific iCBT treatments are also limited in that they leave little room for clinician and patient preferences. One approach to iCBT aims to address these limitations by combining individually tailored treatment according to the participant's needs and symptoms with transdiagnostic components. The idea is to address specific problems and comorbidities with the aim of increasing the scope of iCBT, and possibly increasing motivation and improving treatment outcome. Two previous trials have been conducted on tailored Internet-based treatment for anxiety disorders with comorbid anxiety and depression [17,18]. In this trial we further developed the protocol to treat symptoms of anxiety and depressive symptoms in the presence of panic attacks.

Panic attacks are common across psychiatric conditions, but far from all persons with panic attacks develop panic disorder [19]. Persons with recurrent panic attacks may, however, have other conditions that they want to address in treatment along with their panic symptoms, such as sleep problems, worry, low mood, and other symptoms. However, no trials have tested iCBT specifically for persons presenting with panic attacks regardless of the presence of a diagnosis such as panic disorder and social anxiety disorder. A few previous trials have been conducted on iCBT for younger persons with anxiety disorders. Two studies have investigated the effects of iCBT for social anxiety disorder in high school students [20] and in university students [21]. While these two studies found effects, adherence to treatment was poor. We therefore decided to target persons with panic symptoms with comorbid anxiety and depressive symptoms in this randomized controlled trial. We examined the effects of individually tailored iCBT for young adults aged 18–30 years and adults aged 31–45 years. We hypothesized that the treatment protocol would lead to reductions in symptom measures of panic, anxiety, and depression. We also expected increased quality of life for the participants in the treatment condition and that the treatment effects would remain at the 12-month follow-up. With regard to the two age groups, we did not expect differential outcomes.

Methods

Participants and Recruitment

We recruited participants via an online list (www.studie.nu) among individuals who had expressed an interest in participating in research on iCBT for panic disorder and generalized anxiety disorder via email. The list consisted of 1459 individuals who had been on the list for a minimum of 12,648 days, maximum 14,652 days, and mean of 13,298 days with a standard deviation of 505 days. They were presented with the project website, which contained information about the trial, how to register, and how to submit written informed consent. Screening consisted of the following questionnaires via the Internet: Montgomery-Åsberg Depression Scale-Self-rated (MADRS-S) [22]; Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) [23]; Beck Anxiety Inventory (BAI) [24]; Quality of Life Inventory (QOLI) [25]; Alcohol Use Disorders Identification Test (AUDIT) [26]; and 13 additional questions with reference to demographic variables. If the participants met the initial inclusion criteria they underwent further screening consisting of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) [27] and the Panic Disorder Severity Scale (PDSS) [28], conducted over the telephone by three clinical psychology MSc students who had completed their clinical training. The outcome measures used in the trial have been shown to have good psychometric properties when administered via the Internet [29], and the same accounts for administering the SCID-I interview over telephone [30]. The principal investigator along with a psychiatrist further checked the SCID-I protocols and PDSS assessments for the participants before they enrolled in the trial. The participants had to have reoccurring panic attacks to be included. They could also fulfill the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, text revision [27] criteria for any specific anxiety disorder, or anxiety disorder not otherwise specified except for obsessive-compulsive disorder and posttraumatic stress disorder, which would lead to exclusion. Participants could also meet the criteria for comorbid major depression, but not as the primary disorder. The participants had to be between the ages of 18 and 30 years or between 31 and 45 years; have a total score of <31 on the MADRS-S and a score <4 points on item 9 (suicidal thoughts) on MADRS-S; not currently be in psychotherapy; if on medication, be on stable dosage for the last 3 months; and not be at risk of alcohol abuse or fulfilling the criteria for current alcohol addiction. As outlined by the CONSORT flowchart (Figure 1), 149 individuals expressed an interest in the trial, which commenced in February 2010. After screening and diagnostic interview, we included 57 participants. For a demographic description of the participants, see Table 1. A separate demographic description for the different age groups is available on request.

Table 1. Demographic description of the participants at pretreatment.

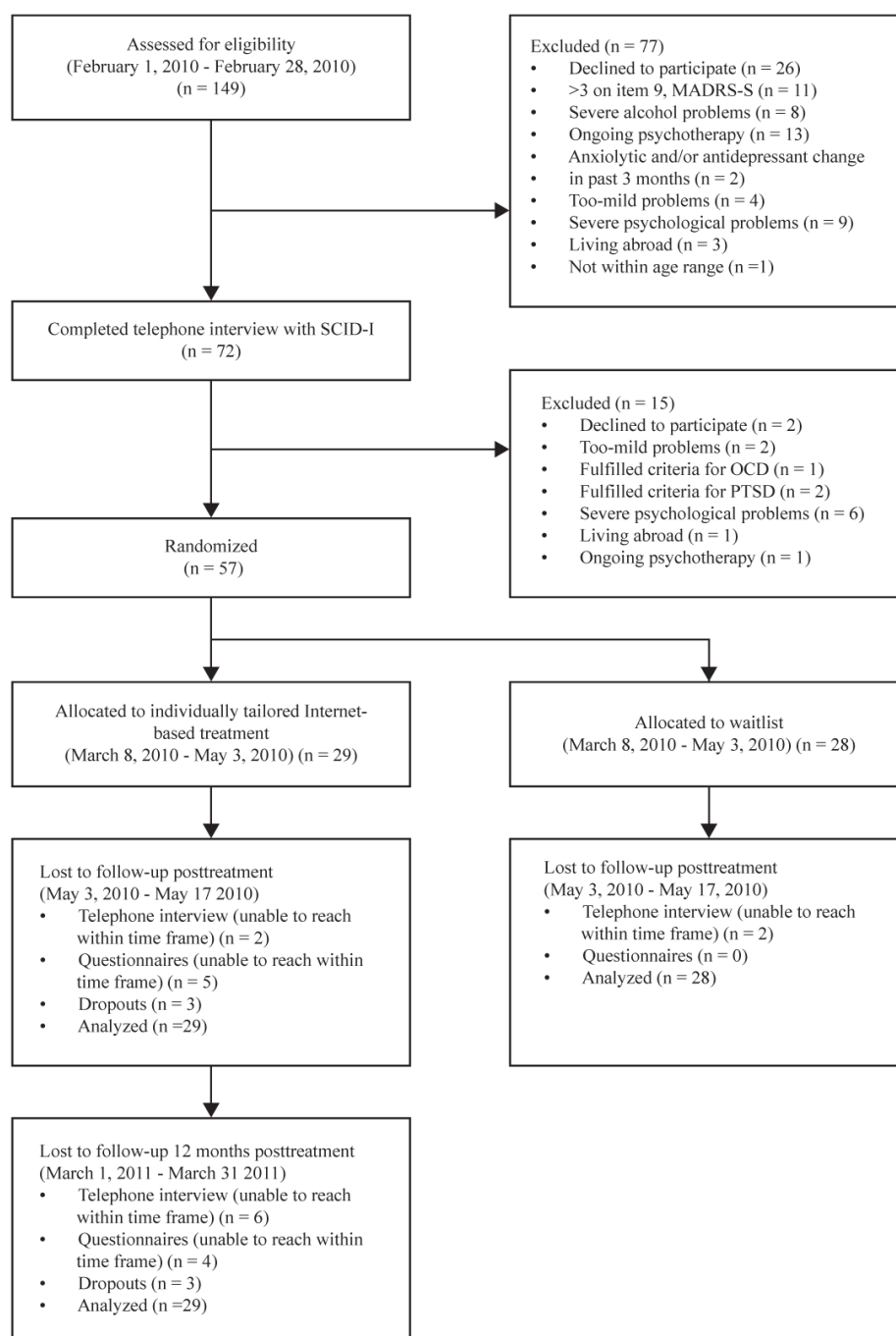
	Treatment group (n = 29)	Control group (n = 28)	Total (n = 57)
Gender, n (%)			
Male	8 (28%)	12 (43%)	20 (35%)
Female	21 (72%)	16 (57%)	37 (65%)
Age (years)			
Mean (SD)	32.3 (7.4)	32.5 (6.5)	32.4 (6.9)
Minimum–maximum	20–45	21–44	20–45
18–30	13 (45%)	12 (43%)	25 (44%)
31–45	16 (55%)	16 (57%)	32 (56%)
Marital status, n (%)			
Single without children	6 (21%)	6 (21%)	12 (21%)
Single with children	1 (3%)	0 (0%)	1 (2%)
Living apart without children	1 (3%)	2 (7%)	3 (5%)
Living apart with children	1 (3%)	1 (4%)	2 (4%)
Married/living together without children	5 (17%)	6 (21%)	11 (19%)
Married/living together with children	15 (52%)	13 (46%)	28 (49%)
Highest educational level, n(%)			
9-year compulsory school	2 (7%)	2 (7%)	4 (7%)
Secondary school	11 (38%)	4 (14%)	15 (26%)
Vocational (completed)	1 (3%)	3 (11%)	4 (7%)
Collage/university (not completed)	5 (17%)	7 (25%)	12 (21%)
Collage/university (completed)	10 (35%)	12 (43%)	22 (39%)
Psychotherapy, n (%)			
No experience	10 (35%)	6 (21%)	16 (28%)
Previous experience	19 (66%)	22 (79%)	41 (72%)
Medication, n (%) ^a			
Ongoing	15 (52%)	12 (43%)	27 (47%)
Completed	4 (14%)	5 (18%)	9 (16%)
No experience	10 (35%)	11 (39%)	21 (37%)
Employment status, n (%)			
Self-employed	2 (7%)	0 (0%)	2 (4%)
Employed	14 (48%)	14 (50%)	28 (49%)
Unemployed	0 (0%)	4 (14%)	4 (7%)
Student	5 (17%)	6 (21%)	11 (19%)
On parental leave	4 (14%)	2 (7%)	6 (11%)
Sick leave	4 (14%)	2 (7%)	6 (11%)
SCID-I diagnosis, n (%) ^b			
Panic disorder	2 (7%)	2 (7%)	4 (7%)
Panic disorder + agoraphobia	24 (83%)	23 (82%)	47 (83%)
Generalized anxiety disorder	2 (7%)	9 (32%)	11 (19%)
Social phobia	1 (3%)	8 (29%)	9 (16%)

	Treatment group (n = 29)	Control group (n = 28)	Total (n = 57)
Anxiety disorder not otherwise specified	1 (3%)	0 (0%)	1 (2%)
Major depression	2 (7%)	3 (11%)	5 (9%)
Comorbid disorders	5 (17%)	13 (46%)	18 (32%)

^a Anxiolytic and/or antidepressant. Needed to be stabilized for 3 months.

^b Structured Clinical Interview for DSM-IV Axis I Disorders.

Figure 1. Flowchart of study participants, point of random allocation, and dropouts at each stage of the trial. MADRS-S = Montgomery-Åsberg Depression Scale; OCD = obsessive-compulsive disorder; PTSD = posttraumatic stress disorder; SCID-I = Structured Clinical Interview for DSM-IV Axis I Disorders.



Treatment

The treatment consisted of 19 CBT modules derived from previous iCBT trials on panic disorder [16], generalized anxiety disorder, social phobia [31], depression [32], and tailored iCBT for anxiety [17,18] and depression (see Table 2 for prescribed modules). In this trial the first module (introduction) and the last module (relapse prevention) were fixed, and the following 17 were optional for the therapists to prescribe: cognitive restructuring (2 modules); panic disorder (2 modules); agoraphobia (1 module); generalized anxiety (3 modules); social anxiety (2 modules); behavioral activation (2 modules); applied relaxation (1 module); stress (1 module); mindfulness (1 module); problem solving (1 module); and insomnia (1 module). The modules are all based on established and evidence-based CBT components. The panic modules, for example, consisted of psychoeducation and interoceptive exposure. All modules included psychoeducation, nearly all contained exposure exercises, and some contained behavioral experiments depending on the content. All modules contained homework assignments for the participants, which consisted of questions on the psychoeducational sections and tasks for the participant to complete, such as exposure exercises. The aim was to prescribe 6–8 modules within an 8-week time frame for each participant. A typical prescription for the participants could be an introduction, cognitive restructuring 1 and 2, panic disorder 1 and 2, agoraphobia, applied relaxation, and relapse prevention. Therapist guidance was included in the trial, since it has been found to improve outcomes when compared with most unguided treatments [33,34]. The therapists were three clinical psychology MSc students who had completed their clinical training and who were supervised by experienced clinical psychologists

(senior authors). The therapists were responsible for 9–10 participants each during the 8 weeks of treatment. The treatment was individually tailored for each participant based on the results of the SCID-I interview and the clinical impression from the telephone interview. The participants were required to have access to a computer with an Internet connection and be able to download the prescribed modules in PDF format through an encrypted contact system, which they also used when communicating with their therapist. The participants were advised to spend 1 week on each prescribed module. If participants had not sent in their homework, which consisted of questions along with worksheets to report on their progress to their therapist within the time frame, their therapist would send out an email reminder. The therapists were instructed to check whether the participants had completed the assignments associated with each module and to provide individual feedback within 24 hours. This was scheduled to occur on Mondays. There were no automatic contacts; all contact was initiated by either the therapist or the participant. If the therapist judged that the participants had completed their homework, the therapist made the next module available through the encrypted contact handling system. Completion of the homework would mean that participants had answered questions about the treatment material and handed in descriptions of their interoceptive exposure, for example. If participants had not completed the previous module, the therapist would give instructions on what needed to be done to be able to move on to the next step. The therapists spent approximately 15 minutes per week per participant (estimated), which included administration as well as reading and responding to emails (estimated 19 emails between therapist and participant throughout the treatment) in the contact handling system.

Table 2. Number of prescribed treatment modules.

Module	Treatment group (n = 29)	Control group (n = 28)	Total (n = 57)
Introduction	29 (100%)	28 (100%)	57 (100%)
Cognitive restructuring 1	26 (90%)	22 (79%)	48 (84%)
Cognitive restructuring 2	26 (90%)	22 (79%)	48 (84%)
Panic 1	29 (100%)	28 (100%)	57 (100%)
Panic 2	29 (100%)	28 (100%)	57 (100%)
Agoraphobia	21 (72%)	22 (79%)	43 (75%)
Social anxiety 1	1 (3%)	3 (11%)	4 (7%)
Social anxiety 2	1 (3%)	3 (11%)	4 (7%)
Generalized anxiety 1	3 (10%)	6 (21%)	9 (16%)
Generalized anxiety 2	3 (10%)	6 (21%)	9 (16%)
Generalized anxiety 3	3 (10%)	6 (21%)	9 (16%)
Behavioral activation 1	0 (0%)	2 (7%)	2 (4%)
Behavioral activation 2	0 (0%)	2 (7%)	2 (4%)
Relaxation (applied)	15 (52%)	15 (54%)	30 (53%)
Sleep management	1 (3%)	1 (4%)	2 (4%)
Mindfulness	3 (10%)	2 (7%)	5 (9%)
Setting boundaries	3 (10%)	1 (4%)	4 (7%)
Solving problems	1 (3%)	0 (0%)	1 (2%)
Preventing relapse	29 (100%)	28 (100%)	57 (100%)

Control Group

The control condition consisted of a waitlist group. Participants were informed that they would receive the treatment after 10 weeks when the treatment group had completed their treatment. They were also informed how to contact the study team if they had any questions during this time frame.

Procedure and Design

We used a randomized controlled design to compare the effects of the treatment against waiting. The participants were divided into two groups so that the two predetermined age groups 18–30 years (young adults) and 31–45 years (adults) were equally represented in each condition. The blocked randomization process was conducted through an online true random number-generation service (random.org) independent of the investigators and therapists. The therapists in this trial were not aware of the participants' age group. The project was approved by the regional ethics committee in Linköping and registered at ClinicalTrials.gov (NCT01296321). We obtained written informed consent through regular mail, which was sent to the study team. At posttreatment participants were instructed via email to complete the follow-up questionnaires and to participate in a semistructured telephone interview carried out by a blinded assessor who had no earlier contact with the participants. The same procedure was repeated at 12 months after treatment completion. We used the PDSS as the primary outcome measure in the trial.

Statistical Analyses

To examine whether the randomization process had succeeded in generating a balanced distribution across the two conditions, we used independent *t* tests and the chi-square test for the demographic data and pretreatment measures. A mixed-models approach with an unstructured covariance structure was endorsed as a way to handle missing data at posttreatment and at the 1-year follow-up. As suggested by Gueorguieva and Krystal [35], we used mixed-effect models due to their advantages over traditional methods of repeated-measures analysis. We calculated between-group and within-group effect sizes (Cohen's *d*) from estimated means and observed pooled standard deviations. We also present a prespecified analysis of a 40% reduction from baseline on PDSS based on observed data, which is equivalent to the participant being much improved, according to the guidelines by Furukawa et al [36].

Results

Treatment Completion

Of the 29 participants in the treatment group, 7 (24%) completed all prescribed modules (6–8) within the 8-week treatment period. A total of 17 participants (59%) completed 50% of the prescribed modules and 14 (48%) completed 75% of the prescribed modules. The mean number of completed modules for the whole group was 5.0 (SD 2.6). The mean number of completed modules in the young adult group was 5.15 (SD 2.34) and the corresponding number in the adult group was 4.19 (SD

3.16). This difference was not statistically significant ($t_{27} = 0.92$, $P = .37$).

Immediate Results: Treatment Versus Control

As is evident in Table 3, the treatment was superior to the control condition with significant interactions on all measures. A mixed-models analysis of the immediate results of the primary outcome measure, PDSS, showed a significant interaction ($F_{1,47.3} = 29.6$, $P < .001$, $d = 1.41$, $d_{\text{young adult}} = 1.59$, $d_{\text{adult}} = 1.20$; see Figure 2). For the secondary outcome measures CORE-OM ($F_{1,46.7} = 10.8$, $P < .01$, $d = 1.01$, $d_{\text{young adult}} = 1.24$, $d_{\text{adult}} = 0.80$),

BAI ($F_{1,49.6} = 4.6$, $P = .04$, $d = 0.57$, $d_{\text{young adult}} = 0.93$, $d_{\text{adult}} = 0.28$), MADRS-S ($F_{1,47.5} = 4.5$, $P = .04$, $d = 0.71$, $d_{\text{young adult}} = 0.88$, $d_{\text{adult}} = 0.60$), and QOLI ($F_{1,48.8} = 5.2$, $P = .03$, $d = 0.54$, $d_{\text{young adult}} = 0.76$, $d_{\text{adult}} = 0.32$), we observed significant interaction effects and moderate to large between-group effect sizes. Age group did not interact with treatment condition across any measure. There was, however, an interaction of time and age group for BAI ($F_{1,49.6} = 7.7$, $P < .01$), which was caused by a slightly lower pretreatment value in the adult control group and thus not reflecting differential treatment effects.

Table 3. Immediate results with intention-to-treat analysis using mixed-effect model estimated means (ES) (n = 57), observed means (OM) (n = 50 for the main outcome measure^a and n = 49 for the secondary outcome measures), and standard deviation (observed) at pre- and posttreatment for the measures of panic, anxiety, depression, and quality of life.

Measure	Time	Group	Age group	ES	OM	SD	95% confidence interval
PDSS ^a	Pre	Treatment	Total	12.54	12.75	4.95	10.62–14.47
			Young adults	12.46	13.17	4.97	9.60–15.32
			Adults	12.63	12.33	5.12	10.05–15.21
		Control	Total	13.71	13.77	5.25	11.74–15.68
			Young adults	13.67	13.67	6.79	10.69–16.65
			Adults	13.75	13.86	3.74	11.17–16.33
	Post	Treatment	Total	6.39	6.54	4.97	4.30–8.48
			Young adults	5.57	6.08	4.03	2.55–8.60
			Adults	7.21	7.00	5.91	4.33–10.09
		Control	Total	13.79	13.81	5.49	11.71–15.85
			Young adults	14.00	14.00	6.58	10.92–17.09
			Adults	13.57	13.64	4.62	10.80–16.33
CORE-OM ^b	Pre	Treatment	Total	1.69	1.69	0.60	1.51–1.87
			Young adults	1.65	1.60	0.61	1.38–1.91
			Adults	1.73	1.78	0.59	1.49–1.97
		Control	Total	1.83	1.83	0.39	1.65–2.02
			Young adults	1.84	1.84	0.43	1.56–2.12
			Adults	1.83	1.82	0.37	1.59–2.07
	Post	Treatment	Total	1.14	1.13	0.56	0.92–1.36
			Young adults	0.96	0.93	0.44	0.65–1.27
			Adults	1.32	1.36	0.60	1.01–1.63
		Control	Total	1.69	1.69	0.53	1.49–1.89
			Young adults	1.65	1.65	0.68	1.35–1.96
			Adults	1.72	1.72	0.40	1.46–1.99
BAI ^c	Pre	Treatment	Total	28.20	28.43	10.06	24.94–31.46
			Young adults	29.15	28.36	11.64	24.31–34.00
			Adults	27.25	28.50	8.63	22.88–31.62
		Control	Total	27.97	27.50	8.91	24.63–31.31
			Young adults	31.25	31.25	10.15	26.21–36.29
			Adults	24.69	24.69	6.92	20.32–29.06
	Post	Treatment	Total	17.89	17.86	8.50	14.16–21.63
			Young adults	15.21	14.82	4.69	9.98–20.45
			Adults	20.58	21.20	10.60	15.25–25.90
		Control	Total	23.03	23.04	9.38	19.65–26.42
			Young adults	23.00	23.00	12.09	17.89–28.12
			Adults	23.06	23.06	7.16	18.63–27.49
MADRS-S ^d	Pre	Treatment	Total	16.24	15.81	7.35	13.38–19.11
			Young adults	15.92	15.00	8.05	11.67–20.18
			Adults	16.56	16.70	6.80	12.73–20.40
		Control	Total	18.13	17.93	8.38	15.20–21.05

Measure	Time	Group	Age group	ES	OM	SD	95% confidence interval
QOLI ^e	Post	Treatment	Young adults	19.50	19.50	7.48	15.07–23.93
			Adults	16.75	16.75	9.06	12.92–20.58
			Total	11.44	11.10	6.65	8.34–14.54
			Young adults	10.57	9.91	4.37	6.14–15.01
			Adults	12.30	12.40	8.57	7.96–16.64
			Total	16.82	16.93	8.47	13.89–19.76
		Control	Young adults	16.08	16.08	8.10	11.65–20.52
			Adults	17.56	17.56	8.94	13.72–21.40
			Total	0.83	0.96	1.92	0.18–1.48
			Young adults	0.93	1.08	2.44	–0.04–1.90
			Adults	0.72	0.83	1.25	–0.15–1.60
			Total	0.74	0.76	1.68	0.08–1.41
	Pre	Treatment	Young adults	0.64	0.64	1.90	–0.37–1.65
			Adults	0.84	0.84	1.55	–0.03–1.72
			Total	1.56	1.67	1.47	0.93–2.20
			Young adults	1.95	2.06	1.44	1.04–2.86
			Adults	1.17	1.24	1.46	0.28–2.05
			Total	0.71	0.70	1.70	0.10–1.31
		Control	Young adults	0.76	0.76	1.69	–0.15–1.68
			Adults	0.65	0.65	1.76	–0.14–1.44

^a Panic Disorder Severity Scale (PDSS).

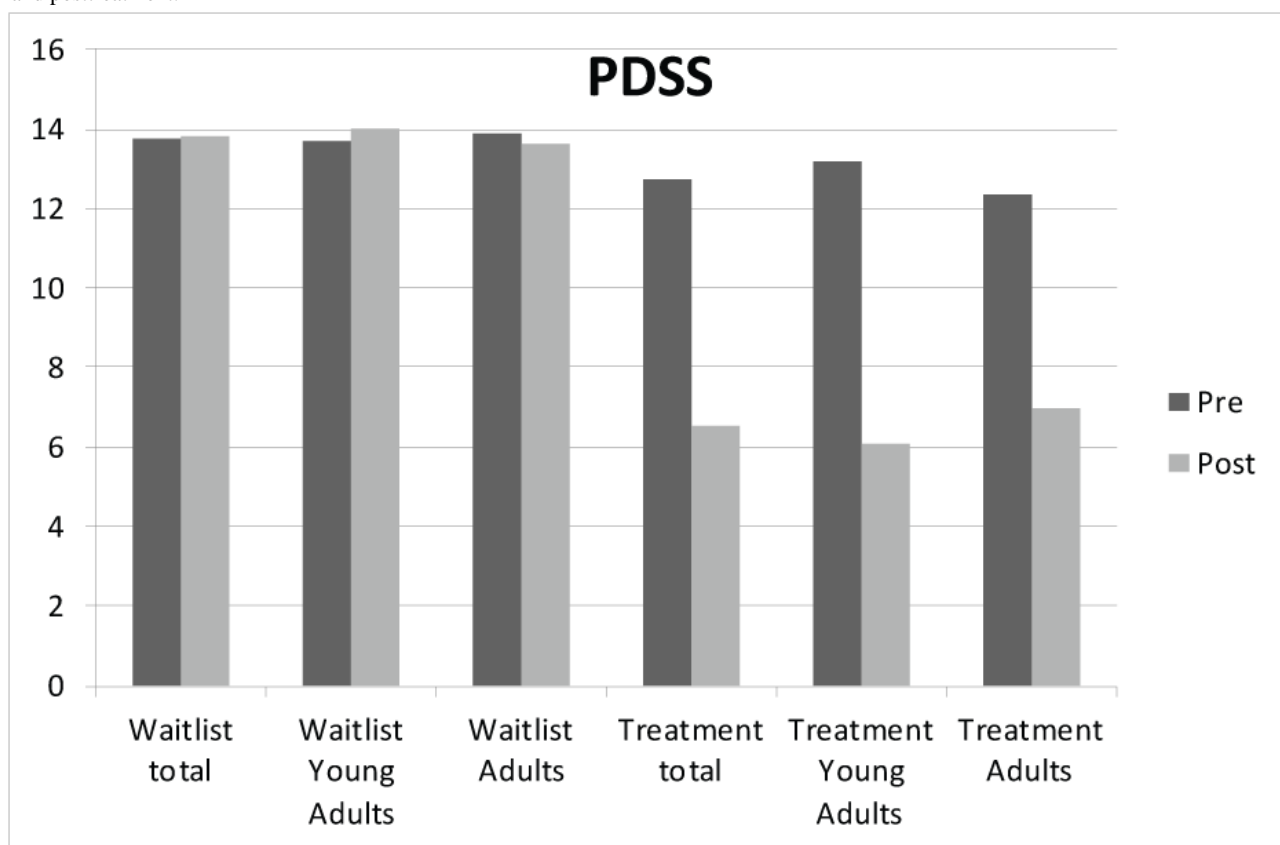
^b Clinical Outcome in Routine Evaluation-Outcome Measure.

^c Beck Anxiety Inventory.

^d Montgomery-Åsberg Depression Rating Scale-Self-rated.

^e Quality of Life Inventory.

Figure 2. Change in observed mean scores in the Panic Disorder Severity Scale (PDSS) for the treatment and waitlist conditions and the two age groups pre- and posttreatment.



Treatment Effects at 1-Year Follow-up

Participants in the treatment condition were contacted 1 year posttreatment and asked to complete the measures again in order to evaluate the long-term effects. A mixed-models analysis showed significant time effects for the primary outcome measure PDSS ($F_{1,18.3} = 19.5$, $P < .001$, estimated mean 4.80, SE 1.18, SD 4.66, $d = 1.66$). We also found effects for the secondary outcome measures BAI ($F_{1,21.2} = 15.2$, $P < .001$, estimated mean 14.45, SE 2.02, SD 9.51, $d = 1.45$), MADRS-S ($F_{1,20.9} = 6.4$, $P < .01$, estimated mean 10.19, SE 1.97, SD 9.53, $d = 0.75$), and QOLI ($F_{1,20.3} = 5.7$, $P < .01$, estimated mean 1.94, SE 0.34, SD 1.55, $d = 0.74$). For these measures, the effect sizes were moderate to large. The exception was the CORE-OM ($F_{1,23.9} = 9.7$, $P < .001$, estimated mean 1.65, SE 0.10, SD 0.45, $d = 0.07$). There was no effect of age group for any of the measures.

Clinical Significance

In the treatment group, 16 of the 24 (67%) participants who took part in the posttreatment telephone interview fulfilled the criteria of a 40% improvement on PDSS. The corresponding number in the control condition was 3 of 27 (11%) participants. Chi-square test showed a significant difference between the two conditions ($\chi^2_1 = 16.1$, $P < .001$). At the 12-month follow-up, 14 of the 20 (70.0%) treated participants who took part in the telephone interview fulfilled the criteria of a 40% improvement on PDSS. Moreover, 16 of the 20 (80.0%) 20 participants no longer met the diagnostic criteria for the diagnosis they had at pretreatment.

Discussion

The aim of this study was to investigate the effects of individually tailored iCBT for panic symptoms along with comorbid anxiety and depressive symptoms on young adults and adults. We found significant treatment effects for all dependent measures immediately following treatment and significant time effects at the 12-month follow-up, showing that a majority of the participants remained stable after completing their treatment. The between-group effect size on the primary outcome measure PDSS was $d = 1.41$ at posttreatment, and the within-group effect size was $d = 1.66$ at the 12-month follow-up. The results of this trial, with moderate to large effect sizes, are consistent with previous trials of iCBT for panic disorder [16] and transdiagnostic iCBT treatment for anxiety [37], which shares some features with our treatment approach. The results of this trial, although focused on individuals with reoccurring panic attacks, are hence in line with previous trials of tailored iCBT for anxiety [17,18]. The participants did not have to meet all the diagnostic criteria for panic disorder to be included in the trial. However, most of them did, with 83% fulfilling the criteria for panic disorder with agoraphobia. This can be regarded as a limitation, as we aimed to target individuals with panic symptoms and not only persons with panic disorder. In light of the sample we recruited, we could expect results similar to those with iCBT for panic disorder. Regarding comorbid conditions, 32% had any comorbid disorder, which could have led to exclusion in previous trials on panic disorder. Previous iCBT trials have been criticized for setting too-strict inclusion criteria, thereby excluding participants with agoraphobia and

comorbid disorders [38]. In this study we found that the treatment may work with less-stringent inclusion criteria.

We found no interaction effect between age group and the treatment condition over all measures; there was, however, an interaction between time and age group for the BAI. While this could indicate that the young adults as a group are more likely to improve spontaneously in the short time frame, it is more probable that this interaction is a chance finding with somewhat lower scores in the adult control group. The effect sizes across all measures showed a tendency for larger effects among the young adults but, due to the small sample size, this requires further investigation for any conclusions to be drawn. Perhaps most important in light of the problems with adherence reported in two previous iCBT trials on social anxiety disorder in high school students [20] and in university students [21] is that adherence was high in this trial, with no difference between the age groups.

This study has limitations. First, the prescription of treatment modules in the study may be unreliable because it was based on a structured diagnostic procedure (SCID-I) and clinical impression by relatively inexperienced clinicians. A more comprehensive clinical assessment, such as a functional analysis, may have resulted in the prescription of other modules. Worth mentioning is also that the treatment modules consisted of modules derived from diagnosis-specific trials and were mainly structured after each diagnosis (eg, panic disorder and not panic symptoms). We aim to address this in future trials on individually tailored treatment for anxiety symptoms, making

the modules symptom specific and less difficult, or less specific for particular diagnoses. Second, the use of a waitlist control condition is a limitation of the present trial. The use of a passive control group means that the effect of nonspecific factors cannot be determined. The lack of a comparison group at the 12-month follow-up makes it impossible to conclude that the improvements over the follow-up period were due only to the effects of the individually tailored treatment. A third limitation is the size of the study, which in turn affects the generalizability of the results. The fourth limitation is that the participants in this study had expressed an interest in iCBT for their problems and therefore the participants may have been highly motivated to undergo treatment. However, we did not measure how highly motivated the participants were. As a fifth limitation, the study was underpowered to detect age differences, and it is possible that small but clinically meaningful differences would have been detected with a larger sample. We do, however, note various clinical implications of the findings in the present trial. Individually tailored iCBT can broaden the use of iCBT to address comorbidity in anxiety disorder, allow for more flexibility, and emphasize the role of the clinician.

Conclusions

The tentative conclusion drawn from these results is that tailoring iCBT can be a feasible approach in the treatment of panic symptoms, comorbid anxiety, and depressive symptoms. Future trials should directly compare individually tailored iCBT for anxiety versus traditional face-to-face CBT, preferably in a clinical setting with larger samples, along with examining how different age groups respond to this treatment format.

Acknowledgments

This trial was funded by a grant from the Swedish Council for Working Life and Social Research. The authors would like to thank the participants in this trial for their involvement.

Conflicts of Interest

None declared.

References

1. Andersson G. Using the Internet to provide cognitive behaviour therapy. *Behav Res Ther* 2009 Mar;47(3):175-180. [doi: [10.1016/j.brat.2009.01.010](https://doi.org/10.1016/j.brat.2009.01.010)] [Medline: [19230862](https://pubmed.ncbi.nlm.nih.gov/19230862/)]
2. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
3. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009 Dec;38(4):196-205. [doi: [10.1080/16506070903318960](https://doi.org/10.1080/16506070903318960)] [Medline: [20183695](https://pubmed.ncbi.nlm.nih.gov/20183695/)]
4. Carlbring P, Westling B, Ljungstrand P, Ekselius L, Andersson G. Treatment of panic disorder via the Internet: a randomized trial of a self-help program. *Behav Ther* 2001;32(4):751-764. [doi: [10.1016/S0005-7894\(01\)80019-8](https://doi.org/10.1016/S0005-7894(01)80019-8)]
5. Klein B, Richards JC, Austin DW. Efficacy of internet therapy for panic disorder. *J Behav Ther Exp Psychiatry* 2006 Sep;37(3):213-238. [doi: [10.1016/j.jbtep.2005.07.001](https://doi.org/10.1016/j.jbtep.2005.07.001)] [Medline: [16126161](https://pubmed.ncbi.nlm.nih.gov/16126161/)]
6. Robinson E, Titov N, Andrews G, McIntyre K, Schwencke G, Solley K. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010;5(6):e10942 [FREE Full text] [doi: [10.1371/journal.pone.0010942](https://doi.org/10.1371/journal.pone.0010942)] [Medline: [20532167](https://pubmed.ncbi.nlm.nih.gov/20532167/)]
7. Paxling B, Almlöv J, Dahlin M, Carlbring P, Breitholtz E, Eriksson T, et al. Guided internet-delivered cognitive behavior therapy for generalized anxiety disorder: a randomized controlled trial. *Cogn Behav Ther* 2011 Sep;40(3):159-173. [doi: [10.1080/16506073.2011.576699](https://doi.org/10.1080/16506073.2011.576699)] [Medline: [21770848](https://pubmed.ncbi.nlm.nih.gov/21770848/)]

8. Furmark T, Carlbring P, Hedman E, Sonnenstein A, Clevberger P, Bohman B, et al. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. *Br J Psychiatry* 2009 Nov;195(5):440-447 [FREE Full text] [doi: [10.1192/bjp.bp.108.060996](https://doi.org/10.1192/bjp.bp.108.060996)] [Medline: [19880935](#)]
9. Berger T, Hohl E, Caspar F. Internet-based treatment for social phobia: a randomized controlled trial. *J Clin Psychol* 2009 Oct;65(10):1021-1035. [doi: [10.1002/jclp.20603](https://doi.org/10.1002/jclp.20603)] [Medline: [19437505](#)]
10. Titov N, Andrews G, Schwencke G, Drobny J, Einstein D. Shyness 1: distance treatment of social phobia over the Internet. *Aust N Z J Psychiatry* 2008 Jul;42(7):585-594. [doi: [10.1080/00048670802119762](https://doi.org/10.1080/00048670802119762)] [Medline: [18612862](#)]
11. Lange A, Rietdijk D, Hudcovicova M, van de Ven JP, Schrieken B, Emmelkamp PM. Intertapy: a controlled randomized trial of the standardized treatment of posttraumatic stress through the internet. *J Consult Clin Psychol* 2003 Oct;71(5):901-909. [doi: [10.1037/0022-006X.71.5.901](https://doi.org/10.1037/0022-006X.71.5.901)] [Medline: [14516238](#)]
12. Andersson G, Waara J, Jonsson U, Malmaeus F, Carlbring P, Ost LG. Internet-based self-help versus one-session exposure in the treatment of spider phobia: a randomized controlled trial. *Cogn Behav Ther* 2009 Jun;38(2):114-120. [doi: [10.1080/16506070902931326](https://doi.org/10.1080/16506070902931326)] [Medline: [20183690](#)]
13. 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](https://doi.org/10.1080/16506070302316)] [Medline: [16291546](#)]
14. Ljótsson B, Hedman E, Andersson E, Hesser H, Lindfors P, Hursti T, et al. Internet-delivered exposure-based treatment vs. stress management for irritable bowel syndrome: a randomized trial. *Am J Gastroenterol* 2011 Aug;106(8):1481-1491. [doi: [10.1038/ajg.2011.139](https://doi.org/10.1038/ajg.2011.139)] [Medline: [21537360](#)]
15. Titov N, Gibson M, Andrews G, McEvoy P. Internet treatment for social phobia reduces comorbidity. *Aust N Z J Psychiatry* 2009 Aug;43(8):754-759. [doi: [10.1080/00048670903001992](https://doi.org/10.1080/00048670903001992)] [Medline: [19629797](#)]
16. Carlbring P, Bohman S, Brunt S, Buhrman M, Westling BE, Ekselius L, et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. *Am J Psychiatry* 2006 Dec;163(12):2119-2125. [doi: [10.1176/appi.ajp.163.12.2119](https://doi.org/10.1176/appi.ajp.163.12.2119)] [Medline: [17151163](#)]
17. Carlbring P, Maurin L, Törngren C, Linna E, Eriksson T, Sparthar E, et al. Individually-tailored, Internet-based treatment for anxiety disorders: A randomized controlled trial. *Behav Res Ther* 2011 Jan;49(1):18-24. [doi: [10.1016/j.brat.2010.10.002](https://doi.org/10.1016/j.brat.2010.10.002)] [Medline: [21047620](#)]
18. 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 Mar;40(1):57-64. [doi: [10.1080/16506073.2010.529457](https://doi.org/10.1080/16506073.2010.529457)] [Medline: [21337215](#)]
19. Goodwin RD. The prevalence of panic attacks in the United States: 1980 to 1995. *J Clin Epidemiol* 2003 Sep;56(9):914-916. [Medline: [14505778](#)]
20. Tillfors M, Andersson G, Ekselius L, Furmark T, Lewenhaupt S, Karlsson A, et al. A randomized trial of Internet delivered treatment for social anxiety disorder in high school students. *Cogn Behav Ther* 2011;40(2):147-157. [doi: [10.1080/16506073.2011.555486](https://doi.org/10.1080/16506073.2011.555486)]
21. Tillfors M, Carlbring P, Furmark T, Lewenhaupt S, Spak M, Eriksson A, et al. Treating university students with social phobia and public speaking fears: Internet delivered self-help with or without live group exposure sessions. *Depress Anxiety* 2008;25(8):708-717. [doi: [10.1002/da.20416](https://doi.org/10.1002/da.20416)] [Medline: [18729147](#)]
22. Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand* 1994 Jan;89(1):21-28. [Medline: [8140903](#)]
23. Barkham M, Margison F, Leach C, Lucock M, Mellor-Clark J, Evans C, et al. Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in the psychological therapies. *Clinical Outcomes in Routine Evaluation-Outcome Measures*. *J Consult Clin Psychol* 2001 Apr;69(2):184-196. [Medline: [11393596](#)]
24. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988 Dec;56(6):893-897. [Medline: [3204199](#)]
25. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the quality of life inventory: a measure of life satisfaction for use in treatment planning and outcome assessment. *Psychol Assess* 1992;4(1):92-101. [doi: [10.1037/1040-3590.4.1.92](https://doi.org/10.1037/1040-3590.4.1.92)]
26. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. *Addiction* 1993 Jun;88(6):791-804. [Medline: [8329970](#)]
27. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th edition, text revision. Washington, DC: American Psychiatric Association; 2000.
28. Shear MK, Brown TA, Barlow DH, Money R, Sholomskas DE, Woods SW, et al. Multicenter collaborative panic disorder severity scale. *Am J Psychiatry* 1997 Nov;154(11):1571-1575. [Medline: [9356566](#)]
29. Carlbring P, Brunt S, Bohman S, Austin D, Richards JC, Öst LG, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Comput Hum Behav* 2007;23(3):1421-1434. [doi: [10.1016/j.chb.2005.05.002](https://doi.org/10.1016/j.chb.2005.05.002)]
30. Rohde P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and II disorders. *Am J Psychiatry* 1997 Nov;154(11):1593-1598. [Medline: [9356570](#)]

31. Andersson G, Carlbring P, Holmström A, Sparthán E, Furmark T, Nilsson-Ihrfelt E, et al. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. *J Consult Clin Psychol* 2006 Aug;74(4):677-686. [doi: [10.1037/0022-006X.74.4.677](https://doi.org/10.1037/0022-006X.74.4.677)] [Medline: [16881775](https://pubmed.ncbi.nlm.nih.gov/16881775/)]
32. Vermark K, Lenndin J, Bjärehed J, Carlsson M, Karlsson J, Öberg 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](https://doi.org/10.1016/j.brat.2010.01.005)] [Medline: [20152960](https://pubmed.ncbi.nlm.nih.gov/20152960/)]
33. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost? *Expert Rev Pharmacoecon Outcomes Res* 2007 Jun;7(3):291-297. [doi: [10.1586/14737167.7.3.291](https://doi.org/10.1586/14737167.7.3.291)] [Medline: [20528315](https://pubmed.ncbi.nlm.nih.gov/20528315/)]
34. Spek V, Cuijpers P, Nyklíček 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](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]
35. Gueorguieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Arch Gen Psychiatry* 2004 Mar;61(3):310-317. [doi: [10.1001/archpsyc.61.3.310](https://doi.org/10.1001/archpsyc.61.3.310)] [Medline: [14993119](https://pubmed.ncbi.nlm.nih.gov/14993119/)]
36. Furukawa TA, Katherine Shear M, Barlow DH, Gorman JM, Woods SW, Money R, et al. Evidence-based guidelines for interpretation of the Panic Disorder Severity Scale. *Depress Anxiety* 2009;26(10):922-929. [doi: [10.1002/da.20532](https://doi.org/10.1002/da.20532)] [Medline: [19006198](https://pubmed.ncbi.nlm.nih.gov/19006198/)]
37. Titov N, Andrews G, Johnston L, Robinson E, Spence J. Transdiagnostic Internet treatment for anxiety disorders: A randomized controlled trial. *Behav Res Ther* 2010 Sep;48(9):890-899. [doi: [10.1016/j.brat.2010.05.014](https://doi.org/10.1016/j.brat.2010.05.014)] [Medline: [20561606](https://pubmed.ncbi.nlm.nih.gov/20561606/)]
38. Westen D, Novotny CM, Thompson-Brenner H. The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychol Bull* 2004 Jul;130(4):631-663. [doi: [10.1037/0033-2909.130.4.631](https://doi.org/10.1037/0033-2909.130.4.631)] [Medline: [15250817](https://pubmed.ncbi.nlm.nih.gov/15250817/)]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test

BAI: Beck Anxiety Inventory

CORE-OM: Clinical Outcomes in Routine Evaluation-Outcome Measure

iCBT: Internet-based cognitive behavior therapy

MADRS-S: Montgomery-Åsberg Depression Scale-Self-rated

PDSS: Panic Disorder Severity Scale

QOLI: Quality of Life Inventory

SCID-I: Structured Clinical Interview for DSM-IV Axis I Disorders

Edited by R Tait, H Christensen, A Callear; submitted 25.05.11; peer-reviewed by L Palmili, E Mayo-Wilson; comments to author 04.07.11; revised version received 02.08.11; accepted 13.09.11; published 26.06.12.

Please cite as:

Silfvernagel K, Carlbring P, Kåbo J, Edström S, Eriksson J, Månson L, Andersson G

Individually Tailored Internet-Based Treatment for Young Adults and Adults With Panic Attacks: Randomized Controlled Trial

J Med Internet Res 2012;14(3):e65

URL: <http://www.jmir.org/2012/3/e65/>

doi: [10.2196/jmir.1853](https://doi.org/10.2196/jmir.1853)

PMID: [22732098](https://pubmed.ncbi.nlm.nih.gov/22732098/)

©Kristin Silfvernagel, Per Carlbring, Julia Kåbo, Sara Edström, Jenny Eriksson, Lisa Månson, Gerhard Andersson. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Superusers in Social Networks for Smoking Cessation: Analysis of Demographic Characteristics and Posting Behavior From the Canadian Cancer Society's Smokers' Helpline Online and StopSmokingCenter.net

Trevor van Mierlo¹, MScCH, MBA, GEMBA; Sabrina Voci², MA; Sharon Lee³, BSC; Rachel Fournier¹, BSC; Peter Selby^{2,4}, MBBS, MHSc, CCFP

¹Evolution Health Systems Inc., San Francisco, CA, United States

²Nicotine Dependence Clinic, Centre for Addiction and Mental Health, Toronto, ON, Canada

³Canadian Cancer Society, Ontario Division, Toronto, ON, Canada

⁴Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Trevor van Mierlo, MScCH, MBA, GEMBA

Evolution Health Systems Inc.

103 10th Street

San Francisco, CA, 94103

United States

Phone: 1 415 839 8113

Fax: 1 415 354 3475

Email: tvnmierlo@evolutionhs.com

Abstract

Background: Online social networks are popular components of behavior-change websites. Research has identified the participation of certain network members who assume leadership roles by providing support, advice, and direction to other members. In the literature, these individuals have been variously defined as *key players*, *posters*, *active users*, or *caretakers*. Despite their identification, very little research has been conducted on the contributions or demographic characteristics of this population. For this study, we collectively categorized key players, posters, active users, and caretakers as superusers.

Objectives: To analyze data from two large but distinct Web-assisted tobacco interventions (WATI) to help gain insight into superuser demographic characteristics and how they use social networks.

Methods: We extracted cross-sectional data sets containing posting behaviors and demographic characteristics from a free, publicly funded program (the Canadian Cancer Society's Smokers' Helpline Online: SHO), and a free, privately run program (StopSmokingCenter.net: SSC).

Results: Within the reporting period (SHO: June 26, 2008 to October 12, 2010; SSC: May 17, 2007 to October 12, 2010), 21,128 individuals registered for the SHO and 11,418 registered for the SSC. Within the same period, 1670 (7.90%) registrants made at least one post in the SHO social network, and 1627 (14.25%) registrants made at least one post in the SSC social network. SHO and SSC superusers accounted for 0.4% (n = 95) and 1.1% (n = 124) of all registrants, and 5.7% (95/1670) and 7.62% (124/1627) of all social network participants, and contributed to 34.78% (29,422/84,599) and 46.22% (61,820/133,753) of social network content, respectively. Despite vast differences in promotion and group management rules, and contrary to the beliefs of group moderators, there were no statistically significant differences in demographic characteristics between the two superuser groups.

Conclusions: To our knowledge, this is the first study that compared demographic characteristics and posting behavior from two separate eHealth social networks. Despite vast differences in promotional efforts and management styles, both WATI attracted superusers with similar characteristics. As superusers drive network traffic, organizations promoting or supporting WATI should dedicate resources to encourage superuser participation. Further research regarding member dynamics and optimization of social networks for health care purposes is required.

(*J Med Internet Res* 2012;14(3):e66) doi:[10.2196/jmir.1854](https://doi.org/10.2196/jmir.1854)

KEYWORDS

Social networks; moderated support; eHealth; community; smoking cessation; Web-assisted tobacco interventions; WATI

Introduction

Tobacco-related illnesses are the leading cause of death in North America, yet 46 million (21%) Americans [1] and 6 million (18%) Canadians [2] continue to smoke. Despite decades of research and prevention, the World Health Organization predicts that tobacco-related illnesses will cause 10 million deaths per year by 2030 [3].

In traditional treatment for tobacco dependence, extensive evidence proves the effectiveness of behavioral interventions such as brief or intensive advice, individual or group counseling, tailored self-help, and telephone quitlines [4,5]. However, despite their proven effectiveness, evidence also shows that these interventions are vastly underused [6]. What is required are effective interventions that have high reach and are easily accessible, implemented, and maintained [7].

As of June 30, 2010, it is estimated that 29% of the world's population has access to the Internet [8]. Research has shown that increasing numbers of individuals are accessing the medium for general health information and to seek assistance with specific addiction and mental health concerns [9,10]. Based on this uptake, a large number of randomized controlled trials and observational studies have investigated how Internet-based interventions can successfully help individuals with problem drinking [11-13], mood and anxiety disorders [14-17], and other conditions [18,19].

Online Social Support for Tobacco Cessation

Following this trend, Web-assisted tobacco interventions (WATI) are proving to be efficacious [20-22]. A popular component of WATI is a social network, comprising online communities of people with a common interest who use a website to communicate with each other, also commonly known as *support groups*, *user forums*, or *discussion boards*. Although there is some evidence that social networks can potentially enhance effectiveness and adherence to eHealth interventions [23,24], very little research has been conducted on how social networks function or who accesses them. There are no best practices for their implementation or maintenance.

In a 2010 study, Cobb et al [25] identified the consistent contributions of specific members over time in QuitNet, a large online community for smoking cessation. In 2010, Selby et al [24] found that 25% of first posts made within StopSmokingCenter.net (SSC) were from recent quitters who were struggling with their quit attempts. Their study also found that 35.0% of first replies were from members who had quit within the past month, 49.0% were from members who had quit for more than 1 month but less than 1 year, and 6.6% of first responses to new messages were from experienced members who had quit for more than a year [24]. In a 2007 analysis of AlcoholHelpCenter.net, an online social network for problem drinkers, Cunningham et al [26] found that discussions clustered around nodes of one or more active users. Most recently in 2011, Jones et al [27] identified members in a self-harm discussion

forum (SharpTalk) who logged on for much greater times than others and mainly posted in response to other participants.

Social network members who assume leadership roles by providing direct support, advice, and direction are defined by Cobb et al as *key players*, by Selby et al as *posters*, by Cunningham et al as *active users*, and by Jones et al as *caretakers*. For consistency in this study we collectively define key players, posters, active users, or caretakers as *superusers*. The purpose of this observational study was to analyze data from two large but distinct WATI to help gain insight into superuser demographic characteristics and how they use social networks.

The Labyrinth of eHealth Applications Online: Promotion and Adherence

While the potential to help individuals through legitimate and validated eHealth interventions is exciting, the explosive growth of Internet access parallels the ever-increasing number of websites and Internet protocols (or IPs or URLs, commonly referred to as *domain names* or *IP addresses*). The overwhelming number of websites available can be especially confusing for general Internet users when distinguishing between legitimate and nonlegitimate eHealth resources [28]. Because the vast majority of search engines and directories generate revenue by promoting website URLs during general searches, and advertisers can maximize their budgets by using algorithms that deliver appealing campaigns to specific demographics and psychographics, it may become increasingly difficult and expensive for legitimate eHealth resources to attract new users.

One method for legitimate eHealth interventions to attract users is through directories such as Beacon [29], an Australia-based organization that publicizes only eHealth IPs that have undergone randomized controlled trials and pass an independent review from a panel of experts. The challenge for important resources such as Beacon is that they require consistent international funding, promotion, and cross-collaborative support to remain effective. Users may also seek new, innovative technologies that may be missed by such directories, since publishing outcome data from randomized controlled trials lags behind adoption of emerging tools.

An alternative grassroots method that can attract general users is through social networking and the retention of superusers. As they expand the size of a network and facilitate discussion, superusers are valuable assets for eHealth social networks to recruit and retain.

For organizations implementing and managing online social networks, identifying and retaining superusers could contribute to the natural development and growth of website traffic and promote adherence. The challenge is determining how to identify superusers, how to attract them, and how to promote their retention. The first step in this process is to begin to understand who superusers are.

Methods

Setting and Program Description

As an initial step designed to investigate superuser characteristics, this observational study analyzed data from two large social networks for smoking cessation: the Canadian

Cancer Society's Smokers' Helpline Online (SHO) (<http://smokershelpline.ca>; Figure 1, Figure 2) [30], and the SSC (<http://www.stopsmokingcenter.net>; Figure 3, Figure 4) [31]. Both programs are available free of charge and are anonymous. We chose the SHO and the SSC social networks due to their large size, active participation, and considerable operational differences.


Figure 1. SmokersHelpline.ca version 6.0 home page.



Figure 2. SmokersHelpline.ca version 6.0 support group home page.

[Home](#) | [Contact us](#) | [Healthcare providers](#) | [Français](#) | [www.cancer.ca](#)


smokershelpline.ca
1 877 513-5333
CONNECT TO QUIT



[Help By Phone](#) | [Online Quit Program](#) | [Success Stories](#) | [Online Community](#) | [Helpful Resources](#)

Smokers' Helpline has proven tips and tools to help you quit successfully. For free, personalized and non-judgmental support, advice and information connect to quit today.

[MEMBER LOGIN](#)
[Forgot username/password?](#)

[ShareThis](#)

Online community support

- [View support group](#)
- [Coping with cravings](#)
- [Hall of Fame](#)
- [Anniversaries](#)
- [Public pledges](#)
- [Testimonials](#)
- [Inspirational e-mails](#)

Online quit program tools

- [Check your smoking](#)
- [Quit Meter](#)
- [Cravings Diary](#)
- [Download self-help guides](#)

Helpful Resources

- [Local Support](#)
- [Tobacco Facts](#)

Welcome to the *Smokers' Helpline Online Support Group*!

If you're looking for advice, if you have some questions or if you need help getting through the rough spots the *Smokers' Helpline Online Support Group* is for you!

If you'd like to ask a question or participate in discussions we ask that you register for our [program](#). Registration is free and it protects our community from spam. Once you register, you must log in to post. Please review our [User Agreement](#).

Our Support Group is moderated by Quit Specialists, to contact them or to learn more about them click [here](#).




There are currently 3 members [logged in](#) and 31 guests browsing the Support Group. Please welcome our newest members: [Today's Active Discussions](#) [Search Messages](#)

Our Forums	Topics	Posts	Last Post
Introduce Yourself! Share about your life as a smoker and why you want to quit.	1907	17388	★ Day 3 by Tiana, Moderator Oct 30, 2010 @ 8:25 PM
General Discussion Your experiences on the tobacco-free journey.	10424	109406	📄 MOOD SWINGS!!! HELP!! by melkshake Oct 31, 2010 @ 10:22 AM
Quit Methods Exploring options and insights.	416	4827	📄 I caved and I'm disg... by jadezoo Oct 18, 2010 @ 3:39 PM
Coping with Cravings and Withdrawal Symptoms Tips and support.	1747	17914	📄 Longest 24 hrs Ever by Pamela, Moderator Oct 28, 2010 @ 11:36 AM
Avoiding relapse and slipping Planning ahead- skills and strategies.	555	6614	📄 I Feel Good! by Dominique, Moderator Oct 29, 2010 @ 12:00 PM
Daily Pledges and Celebrations! Smoke-Free Days, Anniversaries and Hall of Fame.	1971	26453	📄 Sunday's Pledge by paris Oct 31, 2010 @ 9:56 AM
Enhancing Motivation Positive self-talk, benefits and rewards.	63	434	📄 A Quote for Sunday by lithan44 Oct 30, 2010 @ 10:42 PM
New Quit Buddies Private Messaging System Coming Early November...	1	1	📄 New and Improved Quit Bud... by Marianne, Moderator Sep 10, 2010 @ 3:52 PM

NEW - messages since your last visit on October 31, 2010 (11:06 AM)

[Privacy](#) | [Legal](#) | [User agreement](#) | [Bibliography](#)

2010 Canadian Cancer Society. All rights reserved.

Smokers' Helpline Online is for educational purposes only and is not to replace the advice of your family physician or other health care provider. SHO Version 6.0 is Copyright 2005 - 2009 by [Evolution Health Systems Inc.](#) All rights reserved including related methods and software. All worldwide patent rights reserved.

Figure 3. StopSmokingCenter.net version 7.1 home page.

stop smoking center BETA 7.1

LOGIN SIGNUP

"Educator recently logged in"
JASON - Health Educator

COMMUNITY TOOLS & RESOURCES HELP AREA

My Program Personal Profile Public Settings Cravings Diary Private Messaging (0) Motivational Email Text Messaging Forums

I Quit!

TODAY'S TOP DISCUSSIONS

- Will be back when I can, buddies!**
Hot SiO2
August 04 2010 06:58 PM
- First week of August, Hump day pledge!**
Josie, Health Educator
August 04 2010 05:03 PM
- Tony Here from England, Uk**
BlueMadCityFanTony
August 04 2010 04:57 PM
- Comes to prove that it can be done**
cristinas
August 25 2010 11:03 AM

**Signup today
Reward yourself!**
Sign up for your free program and your personalized quit meter will track your money saved, cigarettes not smoked and life gained since quitting.

Quit Meter
Smoke-Free Days: 365
Cigarettes Not Smoked: 8,917
Amount Saved: \$4,895.00
Life Gained
Days: 29hrs - Mins: 16 - Seconds: 49

UPCOMING RESEARCH
Montreal-2nd National Obesity Summit
PRESENTING | APRIL 28 - MAY 1, 2011
Trevor van Mierlo
CEO, founder of Evolution Health
Presenting:
"The relationships of the psychological influence of food and perceived barriers to lifestyle change to body mass index and to utilization of online weight loss tools"

1 2 3 4 5

COMMUNITY
FORUMS
HALL OF FAME
ANNIVERSARY
PUPBLC PLEDGES
HOW TO HELP FELLOW MEMBERS
TESTIMONIALS

TOOLS AND RESOURCES
CHECK YOUR SMOKING
QUIT METER
EDUCATION
EMERGENCY COPING PLAN

HELP AREA
FAQS
FEEDBACK
TUTORIALS
MORE SUPPORT PROGRAMS

OTHER
ABOUT THIS PROGRAM
PRIVACY POLICY
LEGAL INFORMATION
BIBLIOGRAPHY

stop smoking center
Evolution Health

This program is for educational purposes and is not to replace the advice of your family physician or other health care provider. © 2000-2010 Evolution Health Systems Inc.

Figure 4. StopSmokingCenter.net version 7.1 support group home page.

stop smoking center BETA 7.1

LOGIN SIGNUP

"Educator recently logged in"
JASON - Health Educator

COMMUNITY TOOLS & RESOURCES HELP AREA

My Program Personal Profile Public Settings Cravings Diary Private Messaging (0) Motivational Email Text Messaging Forums

Welcome to the Stop Smoking Center Support Group

If you're looking for advice, if you have some questions or if you need help getting through the rough spots the Stop Smoking Center Support Group is for you! Our support group is moderated by trained [Health Educators](#).

To participate in the forums please register. Registration is free and it protects our community from spam. Once you register, you must log in to post. Please review our [User Agreement](#).

MEMBERS RECENTLY LOGGED IN

63 13 SIGNUP NEW

SEARCH FORUMS

	OUR FORUMS	TOPICS	POSTS	LAST POST
NEW	Mod's Corner Daily thoughts, facts and discussion items. What's your take?	1282	3471	Lingering cigarette smell... By peteg Mar 31, 2011 @ 10:52 PM
NEW	Introduce Yourself! Introduce yourself to the group. Share your smoking history, quit date, and what quit method you're using.	7208	80046	Tired of hating myself By Lainey Mar 31, 2011 @ 5:23 PM
NEW	Coping with Withdrawal Symptoms Discussion and advice on Withdrawal Symptoms.	3592	40782	I'm just about to cave. By Penitent Apr 01, 2011 @ 3:42 AM
	Smoking Cessation and Weight-Gain Support and advice on dealing with Weight-Gain.	917	12778	NEVER Be Discouraged By Roxie451 (Shy One) Mar 17, 2011 @ 10:24 PM
NEW	Coping with Cravings A discussion of triggers and tips related to cravings.	2143	26689	Morning crave!!! By Machiavelli Apr 01, 2011 @ 4:29 AM
NEW	Stats Parade Where quitters post their stats and REWARDS!	10934	138610	2 Months For Eyal!! By Penitent Apr 01, 2011 @ 9:07 AM
NEW	Common Support Questions Inviting members to contribute to the reply	829	11101	What the heck is a social smoker? By peteg Mar 28, 2011 @ 9:46 AM
NEW	Forum for New Users & General Discussion Advice from successful participants who have used the site to quit smoking.	20245	210793	Just a thought By Duffis Mar 31, 2011 @ 12:18 PM
	Chewers and Dippers Corner A forum for those who want to overcome their addiction to chewing tobacco.	176	1851	18 Months - No Posts!!! By Ashley, Health Educator Mar 21, 2011 @ 12:44 PM
	TROUBLESHOOTING	TOPICS	POSTS	LAST POST
	Avoiding relapse and slipping Prevention help for those who are afraid of slipping	1167	19077	This is why you think about smoking... By AlohaKeia Mar 10, 2011 @ 3:57 AM
	NOS FORUMS	TOPICS	POSTS	LAST POST
	Forum pour les nouveaux utilisateurs et discussion générale Conseils de participants qui ont utilisé le site et ont réussi à arrêter de fumer.	68	588	Je ne me sens pas bien By Tiana, Health Educator Dec 06, 2010 @ 11:14 AM

NEW - messages since your last visit on March 25, 2011 (1:25 PM)

COMMUNITY

FORUMS
HALL OF FAME
ANNIVERSARY
PUBLIC PLEDGES
HOW TO HELP FELLOW MEMBERS
TESTIMONIALS

TOOLS AND RESOURCES

CHECK YOUR SMOKING
QUIT METER
EDUCATION
EMERGENCY COPING PLAN

HELP AREA

FAQS
FEEDBACK
TUTORIALS
MORE SUPPORT PROGRAMS

OTHER

ABOUT THIS PROGRAM
PRIVACY POLICY
LEGAL INFORMATION
BIBLIOGRAPHY

stop smoking center

Evolution Health

This program is for educational purposes and is not to replace the advice of your family physician or other health care provider. © 2000-2010 Evolution Health Systems Inc.

The Canadian Cancer Society's Smokers' Helpline Online

As a publicly funded program with an internationally recognized and respected brand, SHO undertakes substantial promotional campaigns extending to Internet, television, radio, and print advertising, attracting many smokers. The SHO social network is also used to promote offline activities, such as The Ontario Driven to Quit Challenge. Also, SHO program health educators are required to rigorously monitor content posted by members. An unlimited number of individuals may register for the SHO program. The SHO is a licensed version of the SSC that is modified and updated by the Canadian Cancer Society.

StopSmokingCenter.net

Contrarily, the SSC is not a recognized brand, has never been advertised or promoted, and is maintained as a social enterprise. The organization that operates the SSC (Evolution Health

System Inc, San Francisco, CA, USA) does not actively promote the program, nor does it optimize search engine rankings or actively seek links from other websites. Therefore, members using the SSC are a self-seeking, naturalistic Internet population. Although users must endorse a user agreement to participate in the SSC social network, unlike in the SHO, health educators control very little content posted by members.

SHO and SSC Social Network History, Functionality, and Health Educator Roles

To actively participate in each social network, registration is required. However, actively participating in the social networks is voluntary; in the study period only 7.90% ($n = 1670$) of SHO registrants and 14.25% ($n = 1627$) of SSC registrants made at least one post in the social network (see [Table 1](#)). Program registrants and lurkers do not have to register to browse member conversations within each social network.

Table 1. Total program registrations and social network participation in Smokers' Helpline Online (SHO) and StopSmokingCenter.net (SSC) Web-assisted tobacco interventions.

Characteristic	SHO	SSC
Study period	June 26, 2008–October 12, 2010	May 17, 2007–October 12, 2010
Duration (days)	839	1245
Number of registrants	21,128	11,418
Members with at least one post in the social network, n (%)	1670 (7.90%)	1627 (14.25%)

Every post made in each support group is published immediately. However, to ensure compliance with program rules and regulations, trained health educators review and approve each post through WebTriage (Evolution Health System Inc, San Francisco, CA, USA), a software application designed to facilitate the rapid approval, editing, or deletion of each post.

SHO health educators are paid employees who also manage the Canadian Cancer Society's telephone quitline. SSC health educators are paid employees of Evolution Health Systems Inc. In addition to training in the use of WebTriage, all health educators receive training in cognitive behavioral therapy, stages of change, motivational interviewing, life coaching, data protection, and user privacy. During high-volume seasons, SSC health educators assist SHO health educators with moderating the SHO social network (WebTriage).

Informal Qualitative In-Person Interviews With Social Network Health Educators

For over 5 years, SHO and SSC social network managers have met together on a bimonthly basis to discuss social network management policies, specific incidences within each support group, moderating strategies designed to increase social network traffic, and the management of superuser populations.

On a quarterly basis, SHO and SSC social network managers meet to review cumulative statistical reports, which include the number of program registrants, number of social network posts, traffic statistics (page views, visits, and unique visitors), and basic demographic data collected at registration.

At the onset of this study, we conducted informal qualitative, in-person interviews with health educators and their management

teams. In these interviews, we asked interviewees to reflect on bimonthly and quarterly meetings, each program's promotional and operational differences, and each network's tone and social environment. In these interviews all health educators felt that each program attracted superusers with heterogeneous characteristics, and that in each social network posting behavior and demographics would be correlated.

However, health educator and social network managers indicated that, based on their years of experience, they could not generalize superusers into a single category, and that three distinct types of superusers existed: (1) superusers who cumulatively authored the greatest number of posts (*posters*), (2) superusers who mainly initiated threads (*thread starters*), and (3) superusers who only communicated with certain members (*clique members*), and that management styles differed for each of these superuser subgroups.

Ethical Considerations

All study participants consented to the use of their anonymous data for research purposes. Data collection procedures adhered to international privacy guidelines [32–34] and were in accordance with the Helsinki Declaration of 1975, as revised in 2008 [35]. As the study was based on unidentifiable archival data, the study was deemed to be exempt from further review.

Participants

Registration to both SHO and SSC is anonymous and free of any commitment or fees. From June 26, 2008 to October 12, 2010 (839 calendar days), 1670 members posted at least once in the SHO social network, and there were 84,599 posts made in 7916 threads. From May 17, 2007 to October 12, 2010 (1245

calendar days), 1627 members posted at least once in the SSC social network, and there were 133,753 posts made in 10,967 threads (see [Table 2](#)). During the study periods both programs

used identical question formats and response options in the analysis of demographic characteristics.

Table 2. Total social network activity in Smokers' Helpline Online (SHO) and StopSmokingCenter.net (SSC) Web-assisted tobacco interventions.

Characteristic	SHO	SSC
Study period	June 26, 2008–October 12, 2010	May 17, 2007–October 12, 2010
Duration (days)	839	1245
Number of members with at least one post in the social network	1670	1627
Number of posts in the social network	84,599	133,753
Number of message threads	7916	10,967
Number of posts in shortest thread	1	1
Number of posts in longest thread	73	87

Data Collection

Users completed a baseline questionnaire upon program registration. Questions were based on the North American Quitline Consortium Minimal Data Set [36]. All data were self-reported. On registration, each user is assigned an anonymous user identification, which becomes his or her primary key. We extracted cross-sectional data sets containing posting behaviors and demographic characteristics from each program's customized Structured Query Language database. Health educators' posts were excluded from analysis. Users' primary keys linked their posting behavior with their demographic data.

Posting Behavior and Social Network Participation

In accordance with observations by Cobb et al [25], Selby et al [24], Cunningham et al [26], and Jones et al [27], analysis of posting behavior in both social networks revealed right-skewed distributions, meaning that, cumulatively, most users posted infrequently, and that superusers were responsible for the majority of posts (see [Figure 5](#) and [Figure 6](#)).

Superuser Definition

In each social network, we combined the top 100 posters (ranked according to their total number of posts), thread starters (ranked according to the total number of threads they started), and clique

members (ranked according to the number of threads they participated in) in a single database ($n = 300$). Duplicate entries were removed, leaving a sample of 219 unique superusers. We conducted a power analysis [37] and estimated a total sample size of 95 SHO and 124 SSC superusers to have a power $>80\%$ to test the hypothesis at the $P < .05$ level of significance.

Data Analysis

At registration, SHO and SSC used identical formats and response options for the following demographic variables: gender, age, cigarettes per day, past quit attempts, cohabitant smokers, years smoked, and past nicotine replacement therapy usage. Both programs collected other demographic data, but question formats or response options differed significantly, and thus we excluded those data from analysis.

Descriptive statistics pertaining to general posting behavior and demographic characteristics of both superuser groups are presented first. Next, we conducted sets of univariate logistic regressions to detect differences in demographic characteristics between SHO and SSC superusers. Finally, Pearson correlations were computed to detect relationships between behavior and demographic characteristics in each superuser subgroup. All analyses were performed using SPSS for Windows version 18.0 (IBM Corporation, Somers, NY, USA). The significance level was set at $P < .05$.

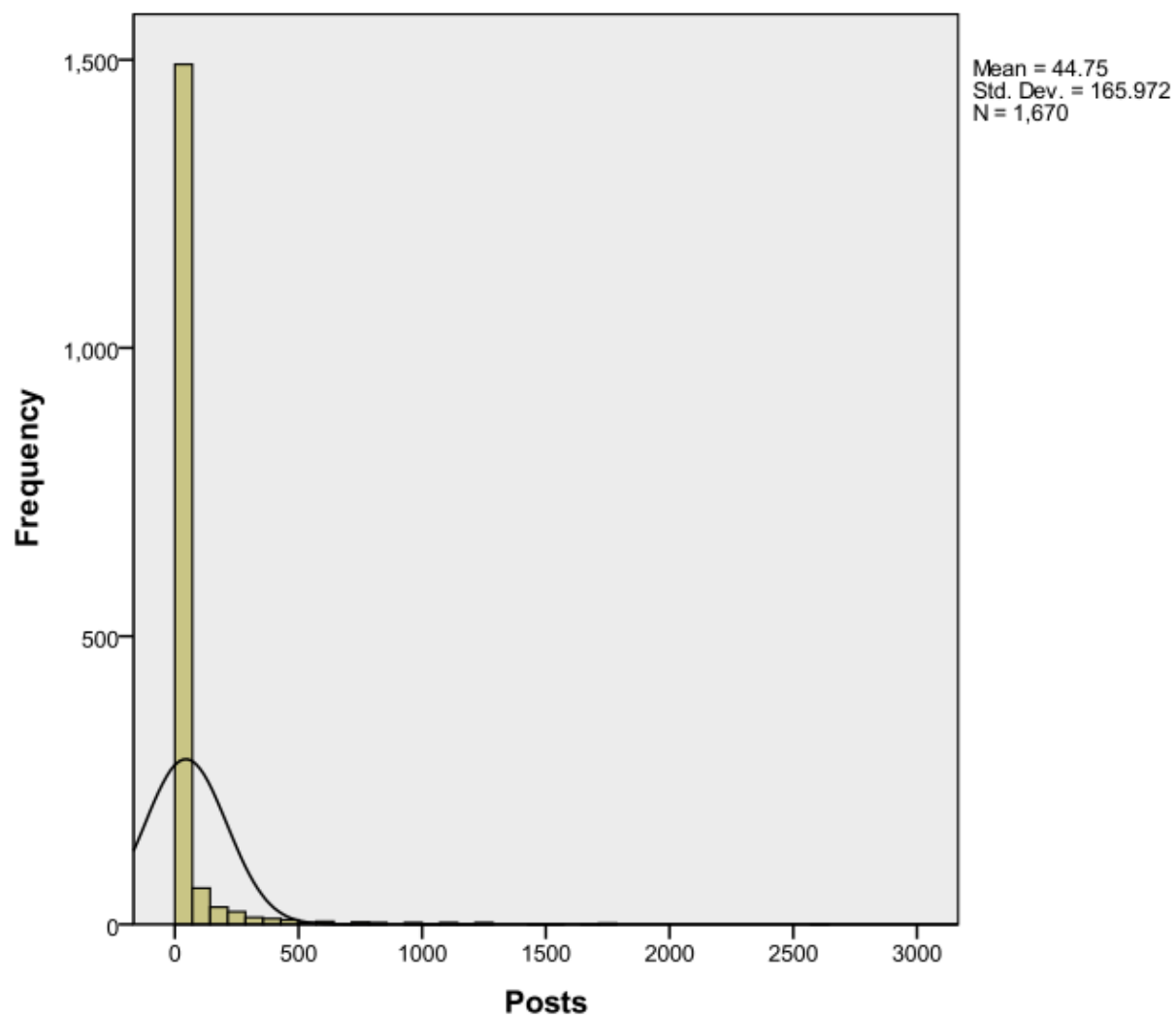
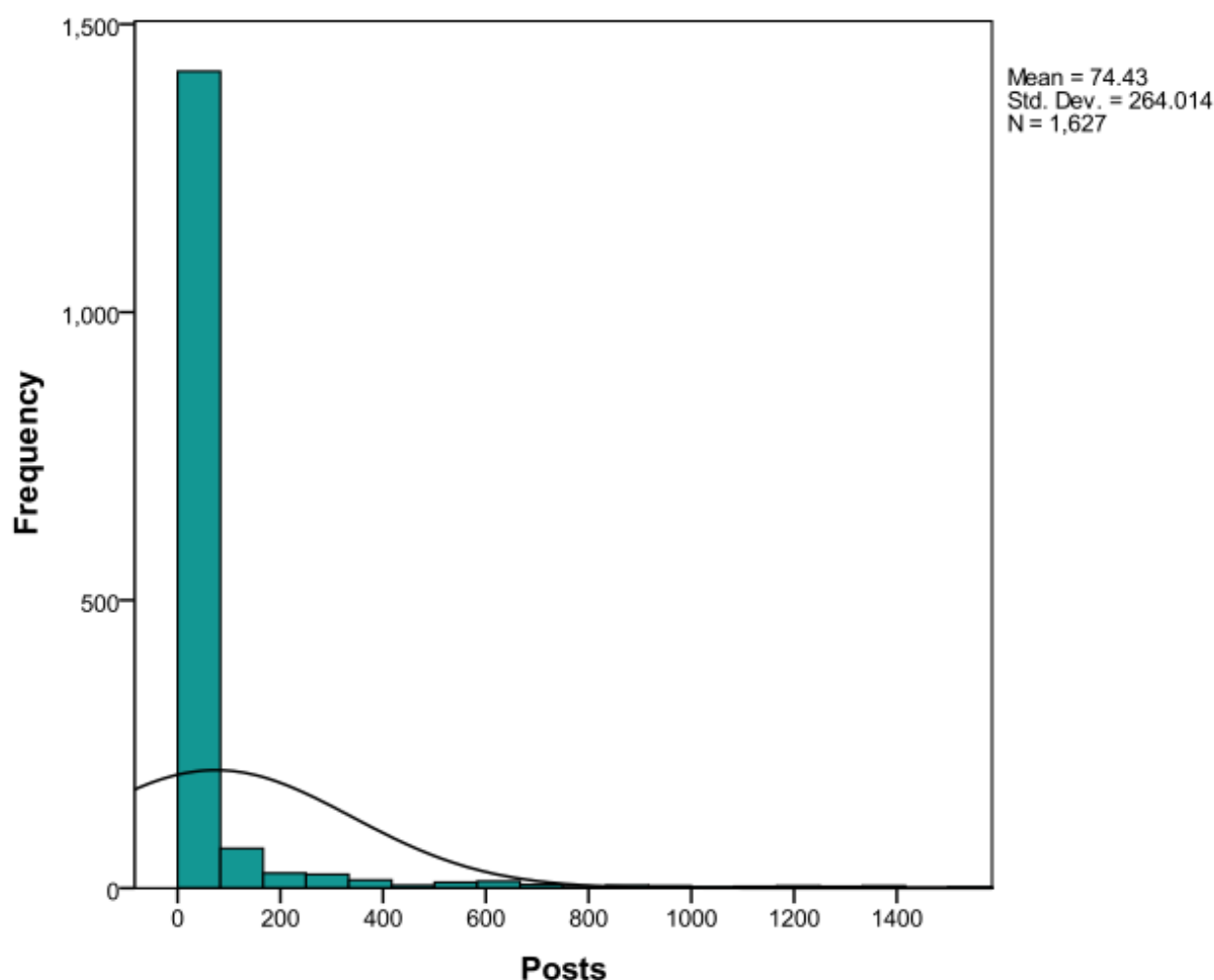
Figure 5. Posting trends in smokershelpline.ca.

Figure 6. Posting trends in StopSmokingCenter.net.

Results

General Posting Behavior and Demographic Characteristics

The 95 SHO and 124 SSC superusers accounted for 0.4% and 1.1% of total program registrants and 5.7% and 7.6% of all

active social network members. They were responsible for 34.78% ($n = 29,422$) and 46.22% ($n = 62,820$) of all social network posts (see [Table 3](#)).

Analysis of demographic characteristics collected at registration (means) and their standard deviations revealed, from the perspective of marketing and moderating techniques, slightly distinct populations (see [Table 4](#)).

Table 3. Superuser social network activity in Smokers' Helpline Online (SHO) and StopSmokingCenter.net (SSC) Web-assisted tobacco interventions.

Characteristic	SHO	SSC
Total number of superusers	95	124
Percentage of total registrants	0.4%	1.1%
Percentage of active social network members	5.7%	7.6%
Total number of posts in social network	84,599	133,753
Posts in social network by superusers, n (%)	29,422 (34.78%)	61,820 (46.22%)

Table 4. Superuser demographic characteristics in Smokers' Helpline Online (SHO) versus StopSmokingCenter.net (SSC) Web-assisted tobacco interventions.

Characteristic	SHO	SSC
Gender (female), n (%)	79 (83%)	91 (73%)
Age (years), mean (SD)	46.0 (11.5)	48.6 (9.5)
Cigarettes smoked per day, mean (SD)	20.3 (8.5)	24.6 (11.2)
Number of past quit attempts, mean (SD)	3.5 (3.2)	3.5 (3.1)
Number of cohabitant smokers, mean (median)	1.1 (1)	0.5 (0)
Smoking duration (years), mean (SD)	17.3 (5.0)	27.0 (10.8)
Past or current nicotine replacement therapy usage, n (%)	50 (52%)	33 (26%)

Comparison of SHO and SSC Superusers

A multivariate logistic regression revealed no statistically significant differences in demographic characteristics between the two populations. Gender, age, cigarettes per day, past quit attempts, cohabitant smokers, years smoked, and past or current use of nicotine replacement therapy were all nonsignificant.

Superuser Subgroup Analysis

To gain further understanding of superuser and superuser-subset characteristics, we determined Pearson correlations to investigate associations between posting behavior and five key demographic

characteristics (age, cigarettes smoked per day, quit attempts, cohabitant smokers, and years smoked).

SHO Pearson Correlations

In the SHO analysis (see Table 5), a statistically significant correlation was revealed between posting behavior and age for all superusers. However, the Pearson correlation was .32, indicating only a small correlation. The only other significant correlation was among clique members, where there was a statistically significant correlation between posting behavior and years smoked. However, this instance also indicated only a small correlation.

Table 5. Bivariate correlations between Smokers' Helpline Online superuser type and demographic characteristics.

Variable	All superusers <i>r</i> (<i>P</i> value)	Posters <i>r</i> (<i>P</i> value)	Thread starters <i>r</i> (<i>P</i> value)	Clique members <i>r</i> (<i>P</i> value)
Age	.32 (.002)	.22 (.08)	.17 (.18)	.23 (.07)
Cigarettes per day	.05 (.66)	.06 (.66)	-.02 (.90)	.06 (.66)
Quit attempts	-.07 (.52)	-.07 (.58)	.03 (.83)	-.06 (.62)
Cohabitant smokers	.03 (.74)	.16 (.21)	.18 (.16)	.20 (.12)
Years smoked	.09 (.37)	.13 (.27)	.14 (.17)	.27 (.03)

SSC Pearson Correlations

In the SSC analysis (see Table 6), there were statistically significant correlations between posting behavior, quit attempts, and years smoked among all superusers, posters, and clique

members. However, these relationships were relatively small (*r* ranging from -.25 to .25). There were no correlations between posting behavior and demographic characteristics among thread starters.

Table 6. Bivariate correlations between StopSmokingCenter.net superuser type and demographic characteristics.

Variable	All superusers <i>r</i> (<i>P</i> value)	Posters <i>r</i> (<i>P</i> value)	Thread starters <i>r</i> (<i>P</i> value)	Clique members <i>r</i> (<i>P</i> value)
Age	.03 (.73)	.07 (.46)	.03 (.76)	.07 (.52)
Cigarettes per day	.12 (.18)	.13 (.21)	.07 (.52)	.18 (.08)
Quit attempts	-.25 (.01)	-.24 (.02)	-.16 (.11)	-.23 (.02)
Cohabitant smokers	-.08 (.40)	-.02 (.90)	.02 (.82)	-.04 (.71)
Years smoked	.24 (.01)	.25 (.01)	.18 (.07)	.25 (.01)

Discussion

At first glance, based on general demographic characteristics outlined in Table 2, SSC superusers appear to have smoked

much longer than SHO superusers, and a smaller number had used or were using nicotine replacement therapy. From these data, marketers and social network health educators may infer that SHO superusers started smoking later in life, and over half have used or are using nicotine replacement therapy.

However, contrary to our original hypothesis, and to information gleaned from informal qualitative in-person interviews with health educators, differences in marketing and the operation of both social networks, tone of discussions, and rules of conduct, superusers in both social networks had similar demographic characteristics.

We were also somewhat surprised to find minimal correlations between posting behavior and demographic characteristics in all eight of the superuser groups analyzed. For example, a Health Canada general population survey found correlations in demographic characteristics and behavior. Although there was little difference in quit rates between Canadian men and women, 29% of smokers aged 20–24 years had quit versus 71% of those aged 45+ years, former male smokers reported an average of 3.2 quit attempts before quitting for good (versus 2.7 quit attempts for females), and at the time of quitting, former smokers reported smoking 18.1 cigarettes per day [38].

Based on the Health Canada general population survey, one might expect to find strong correlations between frequency of posts and age, past quit attempts, or number of cigarettes smoked per day among superusers. Instead, we found only weak relationships between posting behavior, quit attempts, and years smoked, but only in SSC superusers, posters, and clique members.

Strengths and Limitations

To our knowledge, this is the first study that compared demographic characteristics and posting behavior of the most active participants (superusers) from two moderated social networks designed to assist with smoking cessation. A particular strength is that both programs have relatively few barriers to enrollment in comparison with entry barriers that are typical of Internet-based clinical trials [39], as we collected data from WATI operating in their naturalistic Internet environments.

Another strength is the applicability of findings. Results from this study will influence SHO promotional efforts and operations. Web promotion is generally directed at specific and targeted demographics, such as those reported in the Health Canada general population survey. However, to attract superusers, who are not easily defined, promotional efforts should be broadened. Second, as a results of this study management will modify the training programs of SHO and SSC health educators. Health educators will be introduced to strategies that encourage the participation of superusers once identified within the network, and work is underway to develop computer algorithms that will assist health educators with the early identification of superusers and superuser subtypes.

It is important to note that we have used the term superuser and the three superuser subgroups (posters, thread starters, and clique members) only for the purpose of clarifying observations within this specific investigation. These participatory patterns have not been validated, and further research is required to determine

whether they are observable across other types of WATI or health care-based social networks.

To more thoroughly understand superuser characteristics, future studies should compare superuser populations versus nonsuperuser populations (or social network members who create limited posts) and those who register with WATI but do not actively participate in the social network (lurkers). Future research studies should also incorporate the analysis of additional demographic and psychographic characteristics such as occupation, level of education, Goldberg Depression Scale score, Fagerström Test for Nicotine Dependence score, or frequency of Internet usage. These data may be collected at registration; however, in our experience, questions presented to users at registration are most often regarded as intrusive. To maximize use of Web-based programs in a naturalistic setting, registration questions should be kept to an absolute minimum, and other creative means of data collection should be used within interactive program content.

It is also important to note that this study focused only on smokers, and future studies should examine social network behavior and demographic characteristics from superusers, superuser subsets, moderate posters, and lurkers from other condition areas such as depression, panic disorder, problem drinking, self-harm, or healthy weight. This type of research may offer insight into general user characteristics for those who are attracted to Internet programs, and may be used in the development of predictive algorithms.

Summary

The results of this study indicate that superusers play powerful roles within social network traffic. While SHO and SSC superusers accounted for only 0.4% and 1.1% of total program registrants, they were responsible for 34.78% and 46.22% of social network content. The results of this study support the previous research of Cobb et al, Selby et al, Cunningham et al, and Jones et al, indicating that different types of superusers (and superuser subgroups) frequently exist.

Significant and well-recognized barriers to treatment include social stigmas that prevent treatment seeking [40] and inconveniences involved with physically attending treatment or in-person group therapy [41]. Whether individuals act as superusers or passively read the posts from others, the Internet's ability to anonymously and conveniently reach large numbers of individuals can significantly affect health on a population level.

The results of this study suggest that further research in this fast-growing field is required, and that there is potential to maximize the impact of social networks that promote wellness. However, to fully understand the unique mechanisms of Internet-based behavior change, collaboration and knowledge transfer between researchers, nonprofit organizations, and private organizations is recommended.

Acknowledgments

Smokers' Helpline Online (SHO), one component of the Canadian Cancer Society Smokers' Helpline, is sponsored by the Canadian Cancer Society and is available at <http://www.smokershelpline.ca>. StopSmokingCenter.net (SSC) is sponsored by Evolution Health Systems Inc. The authors would like to thank Dr Robert Mann, Senior Scientist at the Centre for Addiction and Mental Health and Associate Professor at the Dalla Lana School of Public Health, University of Toronto, and Dr Walid Hejazi, Professor of Quantitative Reasoning for Management and Associate Professor of International Business at the Rotman School of Management, University of Toronto, for their comments on the study design and an early draft of this paper.

Conflicts of Interest

Trevor van Mierlo is the CEO of Evolution Health Systems Inc. and the owner of StopSmokingCenter.net and other eHealth software platforms. Peter Selby received funds from Schering Canada to provide buprenorphine training, and received honoraria for consultant work, grant funding, advisory board, and/or lectureships from Johnson & Johnson Consumer Health Care Canada; Pfizer Inc, Canada; Sanofi-Synthelabo, Canada; GSK, Canada; Genpharm and Prempharm, Canada; CTI; Evolution Health Systems Inc., Canada; Health Canada; Smoke-Free Ontario; and Canadian Institutes of Health Research. Funding was in compliance with the Canadian Medical Association and the Canadian Psychiatric Association guidelines and recommendations for interaction with the pharmaceutical industry. Sharon Lee is employed by the Canadian Cancer Society. Rachel Fournier is employed by Evolution Health Systems Inc. Sabrina Voci has no interests to declare. None of the authors received any tobacco industry funds.

References

- Centers for Disease Control and Prevention (CDC). Vital signs: current cigarette smoking among adults aged ≥ 18 years --- United States, 2009. *MMWR Morb Mortal Wkly Rep* 2010 Sep 10;59(35):1135-1140 [FREE Full text] [Medline: 20829747]
- Health Canada. 2010 Sep 27. Canadian Tobacco Use Monitoring Survey (CTUMS) 2009. URL: http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/research-recherche/stat/ctums-esutc_2009_graph-eng.php [accessed 2011-05-25] [WebCite Cache ID 5yxK5CsEn]
- World Health Organization. 1999. The World Health Report 1999: Making a Difference URL: http://www.who.int/whr/1999/en/whr99_en.pdf [accessed 2011-09-27] [WebCite Cache ID 6216TNIOz]
- Fiore MD, Jaen CR, Baker TB, Baily WC, Benowitz NL, Currey SJ, et al. US Department of Health and Human Services, Public Health Services. 2008 May. Treating Tobacco Use and Dependence: 2008 Update URL: http://www.surgeongeneral.gov/tobacco/treating_tobacco_use08.pdf [accessed 2011-05-25] [WebCite Cache ID 5yxKZ6uEv]
- Lemmens V, Oenema A, Knut IK, Brug J. Effectiveness of smoking cessation interventions among adults: a systematic review of reviews. *Eur J Cancer Prev* 2008 Nov;17(6):535-544. [Medline: 18941375]
- Orleans CT. Increasing the demand for and use of effective smoking-cessation treatments reaping the full health benefits of tobacco-control science and policy gains--in our lifetime. *Am J Prev Med* 2007 Dec;33(6 Suppl):S340-S348. [doi: 10.1016/j.amepre.2007.09.003] [Medline: 18021909]
- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999 Sep;89(9):1322-1327. [Medline: 10474547]
- Internet World Stats. Miniwatts Marketing Group. 2011 May 21. Internet Usage Statistics: The Internet Big Picture: World Internet Users and Population Stats URL: <http://www.internetworldstats.com/stats.htm> [accessed 2011-05-25] [WebCite Cache ID 5yxKmVjuA]
- Farvolden P, Cunningham JA, Selby P. Using E-health programs to overcome barriers to the effective treatment of mental health and addiction problems. *J Technol Hum Serv* 2009;27(1):5. [doi: 10.1080/15228830802458889]
- Horrrigan J, Rainie L. The Internet's growing role in life's major moments. Washington, DC: Pew Internet & American Life Project; 2006 Apr 19. URL: http://www.pewinternet.org/~media/Files/Reports/2006/PIP_Major%2520Moments_2006.pdf.pdf [accessed 2012-04-10] [WebCite Cache ID 66p2MHZRO]
- Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. *Addiction* 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]
- Doumas DM, Hannah E. Preventing high-risk drinking in youth in the workplace: a web-based normative feedback program. *J Subst Abuse Treat* 2008 Apr;34(3):263-271. [doi: 10.1016/j.jsat.2007.04.006] [Medline: 17600650]
- Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. *Addiction* 2009 Dec;104(12):2023-2032. [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
- Andersson G. Using the Internet to provide cognitive behaviour therapy. *Behav Res Ther* 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
- Farvolden P, McBride C, Bagby RM, Ravitz P. A Web-based screening instrument for depression and anxiety disorders in primary care. *J Med Internet Res* 2003;5(3):e23 [FREE Full text] [doi: 10.2196/jmir.5.3.e23] [Medline: 14517114]
- Griffiths KM, Christensen H. Internet-based mental health programs: a powerful tool in the rural medical kit. *Aust J Rural Health* 2007 Apr;15(2):81-87. [doi: 10.1111/j.1440-1584.2007.00859.x] [Medline: 17441815]

17. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. *J Med Internet Res* 2005;7(1):e7 [FREE Full text] [doi: [10.2196/jmir.7.1.e7](https://doi.org/10.2196/jmir.7.1.e7)] [Medline: [15829479](https://pubmed.ncbi.nlm.nih.gov/15829479/)]
18. Ritterband LM, Cox DJ, Walker LS, Kovatchev B, McKnight L, Patel K, et al. An Internet intervention as adjunctive therapy for pediatric encopresis. *J Consult Clin Psychol* 2003 Oct;71(5):910-917. [doi: [10.1037/0022-006X.71.5.910](https://doi.org/10.1037/0022-006X.71.5.910)] [Medline: [14516239](https://pubmed.ncbi.nlm.nih.gov/14516239/)]
19. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum Internet weight loss program. *J Med Internet Res* 2010;12(1):e9 [FREE Full text] [doi: [10.2196/jmir.1347](https://doi.org/10.2196/jmir.1347)] [Medline: [20350926](https://pubmed.ncbi.nlm.nih.gov/20350926/)]
20. 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](https://doi.org/10.2196/jmir.989)] [Medline: [19000979](https://pubmed.ncbi.nlm.nih.gov/19000979/)]
21. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. *J Med Internet Res* 2008;10(5):e45 [FREE Full text] [doi: [10.2196/jmir.1008](https://doi.org/10.2196/jmir.1008)] [Medline: [19033146](https://pubmed.ncbi.nlm.nih.gov/19033146/)]
22. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. *Arch Intern Med* 2009 May 25;169(10):929-937 [FREE Full text] [doi: [10.1001/archinternmed.2009.109](https://doi.org/10.1001/archinternmed.2009.109)] [Medline: [19468084](https://pubmed.ncbi.nlm.nih.gov/19468084/)]
23. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *J Med Internet Res* 2011;13(1):e30 [FREE Full text] [doi: [10.2196/jmir.1602](https://doi.org/10.2196/jmir.1602)] [Medline: [21393123](https://pubmed.ncbi.nlm.nih.gov/21393123/)]
24. Selby P, van Mierlo T, Voci SC, Parent D, Cunningham JA. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. *J Med Internet Res* 2010;12(3):e34 [FREE Full text] [doi: [10.2196/jmir.1340](https://doi.org/10.2196/jmir.1340)] [Medline: [20719739](https://pubmed.ncbi.nlm.nih.gov/20719739/)]
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. [doi: [10.2105/AJPH.2009.165449](https://doi.org/10.2105/AJPH.2009.165449)] [Medline: [20466971](https://pubmed.ncbi.nlm.nih.gov/20466971/)]
26. Cunningham JA, van Mierlo T, Fournier R. An online support group for problem drinkers: AlcoholHelpCenter.net. *Patient Educ Couns* 2008 Feb;70(2):193-198. [doi: [10.1016/j.pec.2007.10.003](https://doi.org/10.1016/j.pec.2007.10.003)] [Medline: [18022340](https://pubmed.ncbi.nlm.nih.gov/18022340/)]
27. Jones R, Sharkey S, Smithson J, Ford T, Emmens T, Hewis E, et al. Using metrics to describe the participative stances of members within discussion forums. *J Med Internet Res* 2011;13(1):e3 [FREE Full text] [doi: [10.2196/jmir.1591](https://doi.org/10.2196/jmir.1591)] [Medline: [21239373](https://pubmed.ncbi.nlm.nih.gov/21239373/)]
28. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? *J Med Internet Res* 2011;13(2):e35 [FREE Full text] [doi: [10.2196/jmir.1581](https://doi.org/10.2196/jmir.1581)] [Medline: [21531690](https://pubmed.ncbi.nlm.nih.gov/21531690/)]
29. Christensen H, Murray K, Cascar 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](https://pubmed.ncbi.nlm.nih.gov/20528708/)]
30. Canadian Cancer Society. Evolution Health Systems Inc. 2011. Smokers' HelpLine URL: <http://www.smokershelpline.ca/> [accessed 2011-05-25] [WebCite Cache ID 5yxLwr3TZ]
31. Evolution Health Systems Inc. 2011. Stop Smoking Center URL: <http://www.stopsmokingcenter.net/> [accessed 2011-05-25] [WebCite Cache ID 5yxLs1SNn]
32. Office of the Privacy Commissioner of Canada. 2000. Personal Information Protection and Electronic Documents Act: Updated 2006 URL: http://www.priv.gc.ca/legislation/02_06_01_e.cfm [accessed 2012-04-10] [WebCite Cache ID 66pdqU4B9]
33. US Department of Health & Human Services. Health Insurance Portability and Accountability Act of 1996 URL: <http://www.hhs.gov/ocr/office/index.html> [accessed 2011-05-25] [WebCite Cache ID 5yxKziYxy]
34. European Parliament and the Council of the European Union. 2002 Jul 31. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 Concerning the Processing of Personal Data and the Protection of in the Electronic Communications Sector (Directive on Privacy and Electronic Communications) URL: [http://www.aedh.eu/plugins/fckeditor/userfiles/file/Protection%2520des%2520donn%25C3%25A9es%2520personnelles/Directive_EC_2002-58- - eng_.pdf](http://www.aedh.eu/plugins/fckeditor/userfiles/file/Protection%2520des%2520donn%25C3%25A9es%2520personnelles/Directive_EC_2002-58_-_eng_.pdf) [accessed 2012-04-10] [WebCite Cache ID 66pdr3A18]
35. World Medical Association, Inc. 2011. WMA Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects URL: <http://www.wma.net/en/30publications/10policies/b3/index.html> [accessed 2011-05-25] [WebCite Cache ID 5yxLF6EC6]
36. North American Quitline Consortium. Minimal Data Set URL: <http://www.naquitline.org/?page=mds> [accessed 2011-09-27] [WebCite Cache ID 6216lm3Tp]
37. Soper D. Statistics Calculators version 2.0. 2011 May 25. Post-hoc Statistical Power Calculator for Multiple Regression URL: <http://www.danielsoper.com/statcalc/calc09.aspx> [accessed 2011-05-25] [WebCite Cache ID 5yxL9XRtV]
38. Health Canada. 2003. Canadian Tobacco Use Monitoring Survey (CUTM): Quitting Smoking URL: http://www.hc-sc.gc.ca/hc-ps/alt_formats/hecs-sesc/pdf/tobac-tabac/research-recherche/stat/ctums-esutc/fs-if/2003-quit-cess-eng.pdf [accessed 2012-04-10] [WebCite Cache ID 66pdsj5IT]
39. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. *J Med Internet Res* 2009;11(2):e9 [FREE Full text] [doi: [10.2196/jmir.1052](https://doi.org/10.2196/jmir.1052)] [Medline: [19403465](https://pubmed.ncbi.nlm.nih.gov/19403465/)]

40. Cunningham JA, Sobell LC, Sobell MB, Agrawal S, Toneatto T. Barriers to treatment: why alcohol and drug abusers delay or never seek treatment. *Addict Behav* 1993;18(3):347-353. [Medline: [8393611](#)]
41. Collins KA, Westra HA, Dozois DJ, Burns DD. Gaps in accessing treatment for anxiety and depression: challenges for the delivery of care. *Clin Psychol Rev* 2004 Sep;24(5):583-616. [doi: [10.1016/j.cpr.2004.06.001](#)] [Medline: [15325746](#)]

Abbreviations

SHO: Smokers' Helpline Online

SSC: StopSmokingCenter.net

WATI: Web-assisted tobacco interventions

Edited by R Tait, H Christensen, A Caelear; submitted 25.05.11; peer-reviewed by S McIntosh, S Viehbeck; comments to author 07.07.11; revised version received 07.09.11; accepted 28.09.11; published 26.06.12.

Please cite as:

van Mierlo T, Voci S, Lee S, Fournier R, Selby P

Superusers in Social Networks for Smoking Cessation: Analysis of Demographic Characteristics and Posting Behavior From the Canadian Cancer Society's Smokers' Helpline Online and StopSmokingCenter.net
J Med Internet Res 2012;14(3):e66

URL: <http://www.jmir.org/2012/3/e66/>

doi: [10.2196/jmir.1854](#)

PMID: [22732103](#)

©Trevor van Mierlo, Sabrina Voci, Sharon Lee, Rachel Fournier, Peter Selby. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Web-Based Cognitive Behavior Therapy for Depression With and Without Telephone Tracking in a National Helpline: Secondary Outcomes From a Randomized Controlled Trial

Louise Farrer¹, PhD (Clinical Psychology); Helen Christensen¹, PhD; Kathleen M Griffiths¹, PhD; Andrew Mackinnon², PhD

¹Centre for Mental Health Research, The Australian National University, Canberra, ACT, Australia

²Biostatistics Unit, Orygen Youth Health Research Centre, University of Melbourne, Melbourne, VIC, Australia

Corresponding Author:

Louise Farrer, PhD (Clinical Psychology)

Centre for Mental Health Research

The Australian National University

Building 63, Eggleston Rd

Canberra, ACT, 0200

Australia

Phone: 61 (02) 6125 8859

Fax: 61 (02) 6125 0733

Email: louise.farrer@anu.edu.au

Abstract

Background: An earlier report indicated that callers to a telephone counseling service benefited from the addition of an evidence-based Web intervention for depression. It is not known whether the Web intervention would also lower alcohol use and stigma, or improve quality of life and knowledge of depression and its treatments.

Objective: To report the secondary outcomes of a trial of a Web-based cognitive behavior therapy (CBT) intervention for depression, including hazardous alcohol use, quality of life, stigma, depression literacy, and CBT literacy.

Methods: We recruited a sample of 155 callers to Lifeline, a national telephone counseling service in Australia, who met the criteria for moderate to high psychological distress. Participants were randomly assigned to 1 of 4 conditions: (1) Web CBT plus weekly telephone tracking, (2) Web CBT only, (3) weekly telephone tracking only, and (4) neither Web CBT nor telephone tracking. Participants were assessed at preintervention, postintervention, and 6 and 12 months postintervention.

Results: At postintervention, participants who completed the Web intervention either with or without telephone support had lower levels of hazardous alcohol use (without tracking: $P = .008$, effect size = 0.23; with tracking: $P = .003$, effect size = 0.26), improved quality of life (without tracking: $P = .001$, effect size = 0.81; with tracking: $P = .009$, effect size = 0.63), and improved CBT literacy (without tracking: $P = .01$, effect size = 0.71; with tracking: $P < .001$, effect size = 0.80) compared with those who did not receive the Web intervention or telephone support. Results for quality of life and CBT literacy were maintained at 6- and 12-month's follow-up, but differences in hazardous alcohol use were not significantly different between conditions at 6 and 12 months. Although omnibus tests for depression literacy and stigma were nonsignificant, contrasts revealed that those in the Web-only condition showed significantly lower levels of stigma than participants in the control condition at postintervention. This was true for participants in the Web-only and Web plus tracking conditions at 6 months. Similarly, those in the Web-only and Web plus tracking conditions had significantly higher depression literacy at postintervention, and this was maintained in the Web-only condition at 6-months' follow-up. No significant differences were found in depression literacy and stigma between conditions at 12 months.

Conclusions: Evidence-based Web interventions for depression can be effective not only in reducing depression symptoms but also in improving other health outcomes, including quality of life, hazardous alcohol use, and knowledge about effective strategies for depression self-management.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 93903959; <http://www.controlled-trials.com/ISRCTN93903959/> (Archived by WebCite at <http://www.webcitation.org/65y61nSsH>)

(*J Med Internet Res* 2012;14(3):e68) doi:[10.2196/jmir.1859](https://doi.org/10.2196/jmir.1859)

KEYWORDS

eHealth; CBT; cognitive behavior therapy; depression; telephone support

Introduction

Web-based, self-administered cognitive behavior therapy (CBT) programs have been shown to be effective in reducing symptoms of depression [1-4]. Reviews have suggested that the effects of these programs may be enhanced by the provision of guidance from therapists [4]. Based on their recent review of the evidence, Newman and colleagues assert that Web interventions with ongoing assistance from a therapist are superior to other modes of online delivery in the treatment of clinical levels of depression [5]. They argue that Web interventions with less-intensive support (eg, those that involve either no human support or periodic monitoring from a nonspecialist) are more appropriate in the treatment of people with subthreshold mood disorders. However, Titov and colleagues demonstrated that technician (nonspecialist) -administered guidance was as effective as therapist-administered guidance in a Web-based treatment for participants with major depressive disorder [6]. Recently, in a randomized controlled trial of a Web-based CBT program for depression, we found significant reductions in depression symptoms and caseness, both when guidance was provided by a trained volunteer and when no guidance was provided at all, with no significant difference in the effect of the two [7].

While there is evidence for the efficacy of these programs for depression symptoms, less is known about the effects of Web-based programs, guided or not, on secondary outcomes such as alcohol use, quality of life, stigma about depression, or knowledge of depression and its treatment. Accordingly, the current study sought to extend the previously reported outcomes by Farrer and colleagues [7] by examining the effects of the intervention on the following outcomes: hazardous alcohol use, quality of life, stigma, depression literacy, and CBT literacy.

These outcomes are important for several reasons. Depression and hazardous alcohol use frequently co-occur in the general population and in clinical samples [8]. The prevalence of lifetime comorbidity of major depression and alcohol use disorders has been estimated at between 13% and 18% [9,10]. Evidence suggests that depression can increase the likelihood of heavy alcohol use, particularly in women [11]. In recognition of the common causal or etiological mechanisms that may underpin both depressive and alcohol use disorders, CBT is the predominant psychological treatment approach for these conditions. Moreover, computer-based CBT-based interventions have been shown to be effective in the treatment of those with comorbid depression and alcohol use disorders [12].

Quality of life has an evident negative association with depression. Individuals with depression have been shown to have lower levels of functioning and well-being than those with other chronic conditions [13]. Depression is also associated with impaired social and occupational roles [14,15]. Thus, quality of life is an important metric for assessing the impact of an intervention beyond the alleviation of symptoms. It has been argued that a comprehensive assessment of the effectiveness of an intervention should include both the degree

to which the intervention prevents or treats symptoms and the degree to which it enhances the basic elements of well-being [16].

In addition to cognitive restructuring and behavioral activation techniques, comprehensive CBT interventions for depression include a psychoeducational component designed to improve clients' knowledge and understanding of their illness. Education about the signs of depression may improve an individual's ability to detect and respond appropriately to symptoms within themselves and others. There is also evidence that *passive* psychoeducation can have a positive effect on depression symptoms (see [17]). Depression literacy is a term used specifically to describe an individual's awareness of the causes, epidemiology, symptoms, diagnosis, and treatment of depression. Therefore, level of depression literacy is a useful metric for assessing the impact of the psychoeducational component of a CBT-based intervention.

Many individuals with depression may not seek treatment or drop out of treatment prematurely due to the stigma commonly associated with mental health disorders [18]. To reduce the prevalence of depression stigma, educational interventions have been developed to target the common misconceptions held by members of the general public as well as specific groups such as health care professionals [19]. Moreover, individuals with mental disorders have themselves been shown to hold negative views about mental illness [20]. Therefore, educational interventions may effectively reduce stigmatizing attitudes held by those who experience depression. In a previous trial, Web-based psychoeducation was shown to effectively reduce stigma in a community sample with more intense depression symptoms [21], but the effects in a sample with higher needs has not been investigated.

Having established that the Web-based intervention in our recent trial was effective for depression, we hypothesized in this study that Web-based psychoeducation and CBT would be more effective than the control condition in (1) reducing hazardous levels of alcohol use and (2) improving quality of life. Given the psychoeducational and therapeutic content of the intervention, we also expected that the intervention would be effective in (3) reducing stigma relating to depression, (4) improving depression literacy, and (5) improving CBT literacy. We hypothesized that the combination of tracking with the Web intervention would be superior to the other 3 conditions, given the weight of evidence from previous research.

Methods

Data for this study were collected as part of a larger randomized controlled trial of the impact of a Web-based CBT and psychoeducation intervention on depression symptoms [7]. For a full description of the design and methods of the trial, see Farrer et al [7].

Participants

Participants were 155 callers to Lifeline, Australia's 24-hour telephone counseling service, recruited from counseling centers in four major Australian cities between July 2007 and January 2009. A total of 910 callers agreed to be screened by telephone. Of these, 142 (15.6%) were subsequently unable to be contacted, 61 (7%) were later unwilling to participate, and 337 (37%) did not meet eligibility criteria. Respondents were not eligible for inclusion in the trial if they (1) scored less than 22 on the Kessler Psychological Distress Scale [22] (138/337, 41%), (2) had a self-reported diagnosis of schizophrenia or bipolar disorder (89, 26%), or (3) did not have Internet access (67, 20.0%). Ineligible participants were offered brochures sent by mail containing information about the Web intervention used in the trial. Of the 370 people eligible for inclusion in the trial, 155 completed informed consent procedures and preintervention assessments, and were randomly assigned to the trial conditions.

Procedure

Following screening and informed consent procedures, preintervention data were obtained through a self-report questionnaire mailed to participants. We used a block randomization procedure with stratification based on site of recruitment and severity of psychological distress at screening. Allocation of participants to trial conditions was conducted independently by a research assistant not involved in the day-to-day running of the trial. Following randomization, all participants were contacted by telephone and were mailed the relevant materials for their allocated condition.

Intervention and Trial Conditions

The Web-only intervention consisted of Web-based psychoeducation (in week 1 provided by BluePages: bluepages.anu.edu.au) combined with Web-based CBT (in weeks 2–6 provided by MoodGYM: moodgym.anu.edu.au). Both of these Web programs have been shown to reduce depression symptoms in community users [23]. A printed manual containing week-by-week instructions for accessing the Web programs (via a login) was mailed to participants at the start of the trial.

In the Web with tracking condition, participants completed the Web intervention and also received a weekly 10-minute telephone call from a telephone counselor. The purpose of these calls was to address any issues associated with the participants' use of the intervention.

In the tracking-only condition, participants received a weekly 10-minute telephone call from a telephone counselor. These calls focused on various environmental and lifestyle factors associated with depression.

In the control condition, participants received neither the tracking nor Web interventions. Participants in this condition were wait-listed to receive the Web-only intervention following completion of the 6-month follow-up.

Participants in all 4 conditions were free to use the Lifeline telephone counseling service as needed, which provided usual emergency or support services. Any use of this service during the intervention period was additional to the 10-minute

telephone calls offered as part of the Web with tracking and tracking-only intervention conditions. The 10-minute intervention calls were scripted and not intended to provide any form of counseling.

Measures

Data were obtained by self-report questionnaires mailed to participants at preintervention, postintervention (6 weeks following the preintervention assessment), 6 months postintervention, and 12 months postintervention.

Hazardous Alcohol Use

We measured hazardous alcohol use using a 5-item version of the Alcohol Use Disorders Identification Test (AUDIT) [24]. Item scores were summed to provide a total scale score ranging from 0 to 20, with higher scores indicating greater hazardous alcohol use. A score of 5 or above indicates hazardous alcohol consumption [25]. Hazardous alcohol consumption is defined as a pattern of drinking behavior that increases the risk of harmful consequences to the drinker or others [26]. Hazardous drinking behavior does not necessarily reflect the presence of an alcohol use or dependence disorder, although it is a strong risk factor [24]. The internal consistency of the 5-item AUDIT in the current sample was 0.87 ($n = 107$).

Quality of Life

We assessed quality of life using the EUROHIS-QOL 8-item index [27]. The EUROHIS-QOL is composed of 8 items designed to measure the psychological, physical, social, and environmental aspects of quality of life. Items for the EUROHIS-QOL were extracted from the larger World Health Organization Quality of Life (WHOQOL) -100 and WHOQOL-BREF scales, both of which have well-established psychometric properties [28,29]. Item responses were summed, producing a total score ranging from 0 to 32, with higher scores indicating higher quality of life. The internal consistency of the EUROHIS-QOL for the current sample was 0.79 ($n = 151$).

Stigma

We assessed stigma using the personal stigma subscale of the Depression Stigma Scale [21,30]. The personal stigma subscale of the Depression Stigma Scale is composed of 9 items reflecting the participants' personal attitudes toward people with depression. The items were scored on a 5-point scale ranging from strongly agree to strongly disagree, yielding a summed total scale score ranging from 0 to 36, with higher scores indicating higher stigma. In the current sample, the internal consistency of this scale was 0.74 ($n = 151$).

Depression Literacy

We assessed depression literacy using 11 items from a 22-item scale that we had developed for use in a previous community-based Web intervention trial [21]. Participants were provided with 11 statements relating to depression and asked to rate them as either true or false. Example statements are "Reckless and foolhardy behavior is a common sign of depression" (false) and "Antidepressants are addictive" (false). Correct responses received a score of 1, and scores were summed to provide a total depression literacy score ranging from 0 to 11, with higher scores indicating higher depression

literacy. The internal consistency of this scale for the current sample was 0.69 ($n = 151$). This scale was previously shown to be sensitive to the effects of an online depression psychoeducation intervention [21,31].

Cognitive Behavior Therapy Literacy

Knowledge of the principles of CBT was assessed using 8 items from an 18-item scale that we had developed for use in a previous community-based Web intervention trial [23]. The 8 items were statements about CBT that participants were asked to rate as either true or false. Example statements are “I should automatically believe my thoughts because they will more often than not be accurate” (false) and “The statement ‘I’m a stupid idiot’ is an example of labeling” (true). Correct responses received a score of 1, and scores from the 8 items were summed to provide a total CBT literacy score ranging from 0 to 8, with higher scores indicating higher CBT literacy. The internal consistency of this scale for the current sample was 0.69 ($n = 148$). This scale was previously shown to be sensitive to the effects of an online CBT intervention [23].

Statistical Analysis

We analyzed data using SPSS release 18.0.1 for Windows (IBM Corporation, Somers, NY, USA) and Stata 10.1 (StataCorp LP, College Station, TX, USA). A significance level of .05 was used for all outcome variables. We analyzed all outcomes on an intention-to-treat basis. With the exception of alcohol use, all variables had distributions suitable for standard statistical methods. The stigma, quality of life, depression literacy, and CBT literacy variables were analyzed using linear mixed-models repeated-measures analysis of variance, with measurement occasion as a within-groups factor and intervention condition as a between-groups factor. We modeled within-person variation using an unstructured covariance matrix. Mixed modeling allows the use of all available data for each participant (as opposed to substituting missing data with estimated values) [32]. Imputing missing values is potentially problematic, as it can increase the risk of type I error by reducing error variance and artificially increasing degrees of freedom.

The AUDIT had an exponential-shaped distribution, with most participants scoring zero and diminishing numbers with higher scores. Dichotomization at the cut-off for hazardous drinking was not feasible, as it resulted in cells with no variance, having all observations below the cut-off, and low statistical power. Accordingly, we analyzed AUDIT scores using random-intercept

Poisson regression. Poisson regression model counts, in this case, were the total numbers of thresholds (divisions between response categories) crossed on the AUDIT items [33]. The model included intervention and occasion as predictors and an intercept for each participant. The latter random parameters account for within-person correlation of responses over occasions and for overdispersion. Overdispersion refers to observations having greater variance than the mean, as implied by the Poisson model. This analysis was undertaken using the glamm (generalized linear latent and mixed models) procedure in Stata [34].

For each outcome, we used planned contrasts to compare intervention and control groups postintervention, and at 6- and 12-months’ follow-up. We estimated effect sizes by dividing the mean difference between conditions postintervention and at 6- and 12-months’ follow-up by the pooled standard deviation of the groups. Data inspection prior to analysis revealed that two participants in the Web with tracking condition had aberrant patterns of change, each having moderate scores on the Center for Epidemiologic Studies Depression Scale at preintervention and extremely severe scores at postintervention. In one case, we were able to link this to the diagnosis of a life-threatening illness. Due to distortion of the variance and violation of the assumption of normality [35], these participants were removed for the main data analysis. The achieved sample size of 155 participants maintained power above 80% to detect differences between treatment arms of 0.5 SD.

Results

Participation Rates

Of 370 eligible respondents, 155 completed informed consent procedures and preintervention assessments, and were randomly assigned to trial conditions, resulting in a 41.9% acceptance rate. Of the 155 participants, 107 (69.0%) returned postintervention surveys, and 92 (59%) completed the 6-month follow-up. Table 1 shows observed means for outcomes across conditions at preintervention, postintervention, 6-month follow-up, and 12-month follow-up. At preintervention, we found no significant differences between participants in each condition on any of the variables of interest: quality of life ($F_{3,151} = .10$, $P = .96$), hazardous alcohol use ($F_{3,103} = .55$, $P = .65$), stigma ($F_{3,147} = 1.21$, $P = .31$), depression literacy ($F_{3,147} = .98$, $P = .40$), and CBT literacy ($F_{3,144} = 1.41$, $P = .24$).

Table 1. Observed mean scores for hazardous alcohol use, quality of life, stigma, depression literacy, and cognitive behavioral therapy (CBT) literacy for each intervention condition across measurement occasions.

Measure	Condition	Measurement occasion							
		Preintervention		Postintervention		6-month follow-up		12-month follow-up	
		n ^a	Mean (SD)	n ^a	Mean (SD)	n ^a	Mean (SD)	n ^a	Mean (SD)
Hazardous alcohol use									
	Web only	38	4.53 (5.37)	27	2.78 (4.12)	23	3.22 (4.47)	21	3.57 (5.20)
	Web with tracking	45	4.73 (5.76)	17	2.35 (3.10)	20	2.85 (2.70)	14	2.36 (2.37)
	Tracking only	36	3.81 (4.57)	31	3.90 (4.31)	24	3.21 (4.71)	22	3.82 (4.90)
	Control	34	3.09 (4.27)	24	3.92 (5.63)	21	2.57 (3.96)		—
Quality of life									
	Web only	38	12.24 (5.72)	27	16.70 (6.83)	23	17.30 (7.38)	21	17.90 (5.22)
	Web with tracking	45	12.42 (5.22)	17	14.71 (6.75)	20	16.10 (6.27)	14	15.86 (6.86)
	Tracking only	37	12.73 (6.64)	33	14.06 (6.28)	27	14.78 (6.97)	22	15.77 (7.67)
	Control	35	12.06 (4.81)	27	11.56 (5.91)	22	11.14 (6.20)		
Stigma									
	Web only	38	10.26 (5.36)	27	7.37 (5.10)	22	8.09 (5.76)	21	6.57 (4.13)
	Web with tracking	43	11.16 (4.63)	20	10.80 (6.43)	20	9.60 (5.83)	14	10.57 (7.65)
	Tracking only	36	12.61 (6.07)	31	12.42 (5.45)	26	11.58 (5.01)	22	12.09 (5.75)
	Control	34	11.62 (5.61)	27	11.96 (4.68)	22	13.73 (5.39)		
Depression literacy									
	Web only	37	4.27 (2.46)	26	5.96 (2.18)	21	5.57 (2.09)	19	6.47 (2.37)
	Web with tracking	44	4.09 (2.06)	20	5.35 (2.35)	19	5.53 (2.50)	13	5.46 (2.76)
	Tracking only	36	4.67 (2.19)	31	4.42 (2.09)	24	4.33 (2.06)	18	5.00 (1.61)
	Control	34	4.88 (2.33)	27	5.19 (2.80)	20	4.95 (2.67)		
CBT literacy									
	Web only	37	4.35 (1.89)	27	5.67 (1.62)	22	5.77 (1.27)	21	6.33 (1.43)
	Web with tracking	42	4.26 (2.08)	20	5.80 (1.40)	20	5.80 (2.02)	14	5.79 (1.76)
	Tracking only	35	5.06 (1.81)	31	4.10 (1.80)	25	4.48 (1.71)	21	4.67 (1.98)
	Control	34	4.29 (1.80)	27	4.41 (1.95)	20	4.35 (1.98)		

^a Different numbers are due to incomplete responses and unequal dropout between conditions.

Correlations Between Secondary Outcomes and Depression Symptoms at Preintervention, Postintervention, and Follow-ups

A strong negative correlation was found between preintervention depression symptoms and quality of life ($r = -.52$, $P < .001$). We found no other significant correlations: hazardous alcohol use ($r = .04$, $P = .68$), stigma ($r = .15$, $P = .06$), depression literacy ($r = .15$, $P = .07$), and CBT literacy ($r = -.04$, $P = .63$).

At postintervention, depression symptoms and quality of life were negatively correlated ($r = -.65$, $P < .001$), but no other significant correlations were found for hazardous alcohol use ($r = .09$, $P = .38$), stigma ($r = .15$, $P = .13$), depression literacy ($r = .08$, $P = .44$), and CBT literacy ($r = -.15$, $P = .12$).

The pattern of results was similar at the 6- and 12-month follow-ups. Quality of life was negatively correlated with

depression at 6 months ($r = -.77$, $P < .001$) and 12 months ($r = -.69$, $P < .001$). At 12 months, stigma was positively associated with depression symptoms ($r = .29$, $P = .03$). No other significant correlations were found: hazardous alcohol use (6 months: $r = .06$, $P = .58$; 12 months: $r = .15$, $P = .28$), stigma (6 months: $r = .17$, $P = .11$), depression literacy (6 months: $r = .06$, $P = .60$; 12 months: $r = -.07$, $P = .63$), and CBT literacy (6 months: $r = -.18$, $P = .10$; 12 months: $r = -.17$, $P = .20$).

Hazardous Alcohol Use

At preintervention, postintervention, 6-month follow-up, and 12-month follow-up, 49/153 (32%), 22/99 (22%), 17/88 (19%), and 14/57 (25%) participants reported a hazardous level of alcohol use, respectively. There were no significant differences in the proportions of participants reporting hazardous use between conditions at any time point. Table 1 shows mean AUDIT scores for each intervention on each occasion. Planned

comparisons of parameters in the random intercept Poisson regression model showed that postintervention participants who received the Web-only or the Web with tracking interventions had a greater decline than those in the tracking-only ($\chi^2_1 = 4.7$, $P = .03$; $\chi^2_1 = 10.4$, $P < .01$) or control condition ($\chi^2_1 = 5.8$, $P = .02$; $\chi^2_1 = 11.3$, $P < .01$). Differences between the two Web interventions and between the tracking-only and the control condition were not significant ($\chi^2_1 = 0.5$, $P = .50$; $\chi^2_1 = 0.3$, $P = .61$). At the 6-month follow-up, change from preintervention was significantly different for the Web-only and Web with tracking conditions compared with the control group ($\chi^2_1 = 4.0$, $P < .05$; $\chi^2_1 = 5.4$, $P = .02$). There were no significant differences between conditions at the 12-month follow-up.

At postintervention, effect sizes were 0.23 (95% CI –0.32 to 0.79) for the Web-only condition and 0.26 (95% CI –0.39 to 0.91) for the Web with tracking condition, compared with the control condition. Compared with tracking only, effect sizes were 0.27 (95% CI –0.25 to 0.78) for the Web-only condition

and 0.31 (95% CI –0.31 to 0.93) for the Web with tracking condition.

Quality of Life

The occasion-by-condition interaction was also significant for quality of life ($F_{8,96.4} = 2.39$, $P = .02$). At postintervention, participants who received the Web with tracking and Web-only interventions has significantly greater improvements in quality of life than participants in the control condition, but not the tracking-only condition (Table 2). At 6 months, the Web with tracking (contrast estimate = 5.12, 95% CI 1.81–8.45, $P = .003$) and Web-only (contrast estimate = 5.06, 95% CI 1.93–8.18, $P = .002$) conditions remained superior to the control condition. No significant differences were observed between conditions at 12 months.

At postintervention, effect sizes were 0.81 (95% CI 0.25–1.36) for the Web-only condition and 0.63 (95% CI –0.02 to 1.27) for the Web with tracking condition, compared with the control condition. Compared with tracking only, effect sizes were 0.41 (95% CI –0.11 to 0.92) for the Web-only condition and 0.22 (95% CI –0.39 to 0.83) for the Web with tracking condition.

Table 2. Contrast estimates, 95% confidence intervals (CI), test of significance values, and effect size estimates for contrasts of pre- to postintervention change between intervention and control conditions.

Outcome	Statistic	Web only vs control	Web only vs tracking only	Web with tracking vs control	Web with tracking vs tracking only
Quality of life	Contrast estimate	4.20	2.31	3.88	1.98
	95% CI	1.70–6.70	–0.08 to 4.70	0.99–6.76	–0.81 to 4.78
	P value	.001	.06	.009	.16
Stigma	Contrast estimate	–2.19	–.97	–.22	.10
	95% CI	–4.35 to –0.03	–3.05 to 1.11	–2.59 to 2.14	–1.30 to 3.30
	P value	.047	.36	.85	.39
Depression literacy	Contrast estimate	0.61	1.07	1.16	1.61
	95% CI	–0.42 to 1.65	0.07–2.07	0.03–2.29	0.52–2.71
	P value	.24	.04	.045	.004
CBT literacy	Contrast estimate	1.10	1.87	1.79	2.56
	95% CI	0.31–1.88	1.10–2.63	0.94–2.64	1.73–3.39
	P value	.01	<.001	<.001	<.001

Stigma

The interaction of condition and occasion was nonsignificant for stigma ($F_{8,96.5} = 1.73$, $P = .10$). Nonetheless, we found several between-condition contrasts to be significant. At postintervention, participants in the Web-only condition showed significantly lower levels of stigma than participants in the control condition. At 6 months, stigma was significantly reduced in participants in the Web-only condition (contrast estimate = –3.29, 95% CI –5.97 to –0.61, $P = .02$) and the Web with tracking condition (contrast estimate = –2.88, 95% CI –5.71 to

–0.05, $P = .046$), relative to participants in the control condition. No significant differences were found between conditions at the 12-month follow-up.

At postintervention, effect sizes were 0.94 (95% CI 0.38–1.50) for the Web-only condition and 0.17 (95% CI –0.42 to 0.77) for the Web with tracking condition, compared with the control condition. Compared with tracking only, effect sizes were 0.96 (95% CI 0.41–1.50) for the Web-only condition and 0.24 (95% CI –0.34 to 0.82) for the Web with tracking condition.

Depression Literacy

The interaction of condition and occasion was also nonsignificant for depression literacy ($F_{8,96.5} = 1.75$, $P = .10$). However, at postintervention, contrasts revealed that participants who received the Web with tracking intervention had significantly higher levels of depression literacy than participants in the control and tracking-only conditions. Participants in the Web-only condition also had significantly improved depression literacy compared with participants in the tracking-only condition. At 6 months, depression literacy remained higher in participants in the Web with tracking condition than in participants in the control condition (contrast estimate = 1.64, 95% CI 0.30–3.00, $P = .02$). No significant differences were found between conditions at 12 months.

At postintervention, effect sizes were 0.31 (95% CI –0.23 to 0.85) for the Web-only condition and 0.01 (95% CI –0.58 to 0.61) for the Web with tracking condition, compared with the control condition. Compared with tracking only, effect sizes were 0.73 (95% CI 0.19–1.26) for the Web-only condition and 0.37 (95% CI –0.21 to 0.96) for the Web with tracking condition.

CBT Literacy

We found a significant condition-by-occasion interaction for CBT literacy ($F_{8,101.9} = 6.29$, $P < .001$). At postintervention, participants who received both Web interventions had significantly higher CBT literacy scores than participants in the control and tracking-only conditions. At 6 months, these improvements were maintained such that CBT literacy was higher in participants in the Web-only (contrast estimate = 1.24, 95% CI 0.33–2.15, $P = .008$) and Web with tracking (contrast estimate = 1.71, 95% CI 0.76–2.67, $P = .001$) conditions relative to the control condition. CBT literacy was also higher in the Web-only (contrast estimate = 1.48, 95% CI 0.61–2.37, $P = .001$) and Web with tracking (contrast estimate = 1.96, 95% CI 1.03–2.88, $P < .001$) conditions relative to the tracking-only condition. At 12 months, participants in both the Web-only (contrast estimate = 1.93, 95% CI 0.94–2.92, $P < .001$) and Web with tracking (contrast estimate = 1.91, 95% CI 0.82–3.00, $P = .001$) conditions showed higher CBT literacy than participants in the tracking-only condition.

At postintervention, effect sizes were 0.71 (95% CI 0.16–1.26) for the Web-only condition and 0.80 (95% CI 0.18–1.42) for the Web with tracking condition, compared with the control condition. Compared with tracking only, effect sizes were 0.92 (95% CI 0.37–1.46) for the Web-only condition and 1.03 (95% CI 0.41–1.64) for the Web with tracking condition.

Discussion

Analysis of the secondary outcomes of this trial suggests that Web-based CBT is effective in the short-term minimization of hazardous alcohol use, and in the immediate- and longer-term improvement of quality of life and CBT literacy. For these outcomes, weekly telephone tracking provided by a lay telephone counselor did not confer any advantage over delivery of the Web intervention alone. This result is consistent with that obtained in the analysis of the primary outcome of depression for this trial [7], where we found that Web-based CBT without

telephone tracking was as effective as the intervention with telephone tracking. Primary outcome analyses also revealed that telephone tracking was associated with significantly greater dropout in the Web with tracking condition than in other conditions. As might be expected, given that the same participants were involved, dropout was higher in the Web with tracking condition on the secondary outcome measures as well.

At preintervention, average levels of alcohol use approached the cut-off indicating hazardous use. Participants who received the Web intervention significantly reduced their alcohol use, suggesting that Web-based CBT for depression may be effective in the prevention of hazardous alcohol use. CBT-based interventions have been shown to be effective in the treatment of alcohol use disorders. However, little is known about the mechanisms of action that underpin how and why CBT is effective. It has been hypothesized that CBT promotes the acquisition of cognitive and behavioral coping skills that enable individuals to manage the life stress and alcohol cues that maintain excessive drinking [36]. However, a review by Morgenstern and Longabaugh did not find evidence to support the mediating role of coping skills [36]. More recent investigations of computer-based CBT for substance use disorders have revisited this hypothesis, suggesting that it may be the quality, not quantity, of coping skills acquired through CBT that leads to reductions in hazardous alcohol use [37].

The Web intervention was also effective in improving quality of life. We found a strong negative correlation between depression and quality of life at preintervention, and improvement in quality of life may be partly associated with the observed reductions in depression symptoms. However, quality of life reflects the broader construct of well-being that includes variables such as satisfaction with general health, finances, social relationships, and activities of daily living. It is possible that the alleviation of depression symptoms, or the acquisition of cognitive and behavioral strategies, or both, are associated with an enhanced ability to effect change in dissatisfying life circumstances or to adopt a more helpful or realistic view of existing circumstances. In any case, the findings show that the effects of the intervention were broader than those of depression symptoms.

The intervention was not effective overall in improving depression literacy or in reducing stigma, although some contrast effects were found. Interventions that focus on providing information about mental illness have been shown to improve both depression literacy and stigma [19]. The length of the psychoeducational component of the current intervention (1 week) was probably not sufficient to produce meaningful change in knowledge of depression, a factor that may also account for the absence of differential change in depression stigma. A trial involving a larger “dose” of Web-based psychoeducation than the current trial (5 weeks) obtained both significant increases in depression literacy and significant reductions in stigma at postintervention [21]. An additional explanation may be the use of the shortened version of the depression literacy scale in the current trial. The hypothesis that stigma reduction is mediated by an improvement in depression literacy has been examined previously [21]. However, in the current study, the intervention was not found to be effective for either depression literacy or

stigma, violating the conditions required to investigate mediation using the causal-steps approach proposed by Baron and Kenny [38].

The current intervention was effective, however, in improving knowledge of the key principles of CBT, which is expected, given that this was the predominant focus of the intervention. The sustained improvement in CBT knowledge observed at the 6- and 12-month follow-ups suggests that participants engaged with the intervention at a depth that created lasting memory for the therapeutic concepts contained in the program.

Limitations

We acknowledged that greater dropout in the Web with tracking condition may have inflated the effects observed in this condition. It is noteworthy, however, that dropout was not greater in the Web-only condition, so that the finding of differential dropout from the Web with tracking condition does not undermine the findings suggesting that the Web intervention conditions are more effective than either the tracking-only or control conditions.

It is possible that differences in the amount of Lifeline use between conditions throughout the trial may have influenced the findings, particularly if participants in the intervention conditions used the service more. However, those in the active treatment conditions (Web only and Web with tracking) declined to use the service during the intervention period, and were significantly less likely than those in the control condition to use the Lifeline service frequently (calling 3 or more times in a 1-month period) at postintervention.

Further, as noted in our report of the primary outcomes of this trial [7], uptake of the trial and intervention completion rates were lower in our study than in some other trials of Web-based treatments for depression [39]. The reasons for this are unclear. Users of telephone counseling services may experience more stressful events that preclude trial completion or may prefer to manage their own health. The current trial was a true effectiveness trial employing the telecounseling organization's existing volunteer workforce for recruitment and tracking, and hence would be expected to be associated with greater recruitment and adherence issues relative to the more controlled environment of an efficacy study [40].

The current study found evidence of efficacy for the secondary outcomes examined, but we did not investigate the mediating variables through which the intervention was effective. Pre-posttest designs do not fully allow for examination of mediation or causal chains. Studies including more frequent observations are needed to examine such effects, and studies of this type need not be controlled trials. Moreover, in some cases, the relationships between key variables of interest were nonsignificant, and thus we could not conduct mediator analyses.

Conclusion

This is the first study to demonstrate that a Web-based psychoeducation and CBT intervention introduced through a telephone helpline is effective in minimizing hazardous alcohol use, and in improving quality of life and knowledge of CBT. The results highlight the potential for low-cost, low-intensity Web-based interventions to produce change in outcomes beyond those that the intervention is primarily designed to produce.

Acknowledgments

Funding for the trial was provided by an Australian Research Council Linkage Project Grant (LP0667970). LF is supported by an Australian Postgraduate Award Industry scholarship. KG is supported by an NHMRC Fellowship (No. 525413) and HC is supported by an NHMRC Fellowship (No. 525411). Nicole Burgess coordinated and conducted data collection for the trial. Dawn O'Neil was involved in obtaining funding for the trial and coordinating Lifeline Australia's collaboration with the Australian National University. Associate Prof Richard O'Kearney was a member of LF's PhD supervisory panel. Trevor Carlyon was also involved in coordinating Lifeline staff and volunteer involvement in the trial and supported N Burgess in her role.

Conflicts of Interest

HC and KG are authors of the Web-based treatment programs BluePages (bluepages.anu.edu.au) and MoodGYM (moodgym.anu.edu.au).

References

1. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009 Dec;38(4):196-205. [doi: [10.1080/16506070903318960](https://doi.org/10.1080/16506070903318960)] [Medline: [20183695](https://pubmed.ncbi.nlm.nih.gov/20183695/)]
2. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
3. Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. *Med J Aust* 2010 Jun 7;192(11 Suppl):S4-11. [Medline: [20528707](https://pubmed.ncbi.nlm.nih.gov/20528707/)]
4. Spek V, Cuijpers P, Nyklíček 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](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]

5. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? *Clin Psychol Rev* 2011 Feb;31(1):89-103. [doi: [10.1016/j.cpr.2010.09.008](https://doi.org/10.1016/j.cpr.2010.09.008)] [Medline: [21130939](#)]
6. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010;5(6):e10939 [FREE Full text] [doi: [10.1371/journal.pone.0010939](https://doi.org/10.1371/journal.pone.0010939)] [Medline: [20544030](#)]
7. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Internet-based CBT for depression with and without telephone tracking in a national helpline: randomised controlled trial. *PLoS One* 2011;6(11):e28099 [FREE Full text] [doi: [10.1371/journal.pone.0028099](https://doi.org/10.1371/journal.pone.0028099)] [Medline: [22140514](#)]
8. Swendsen JD, Merikangas KR. The comorbidity of depression and substance use disorders. *Clin Psychol Rev* 2000 Mar;20(2):173-189. [Medline: [10721496](#)]
9. Grant BF, Harford TC. Comorbidity between DSM-IV alcohol use disorders and major depression: results of a national survey. *Drug Alcohol Depend* 1995 Oct;39(3):197-206. [Medline: [8556968](#)]
10. Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL, et al. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* 1990 Nov 21;264(19):2511-2518. [Medline: [2232018](#)]
11. Dixit AR, Crum RM. Prospective study of depression and the risk of heavy alcohol use in women. *Am J Psychiatry* 2000 May;157(5):751-758. [Medline: [10784468](#)]
12. Kay-Lambkin FJ, Baker AL, Lewin TJ, Carr VJ. Computer-based psychological treatment for comorbid depression and problematic alcohol and/or cannabis use: a randomized controlled trial of clinical efficacy. *Addiction* 2009 Mar;104(3):378-388. [doi: [10.1111/j.1360-0443.2008.02444.x](https://doi.org/10.1111/j.1360-0443.2008.02444.x)] [Medline: [19207345](#)]
13. Wells KB, Stewart A, Hays RD, Burnam MA, Rogers W, Daniels M, et al. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA* 1989 Aug 18;262(7):914-919. [Medline: [2754791](#)]
14. Hecht H, von Zerssen D, Krieg C, Pössl J, Wittchen HU. Anxiety and depression: comorbidity, psychopathology, and social functioning. *Compr Psychiatry* 1989;30(5):420-433. [Medline: [2791535](#)]
15. Whooley MA, Kiefe CI, Chesney MA, Markovitz JH, Matthews K, Hulley SB, CARDIA Study. Depressive symptoms, unemployment, and loss of income: The CARDIA Study. *Arch Intern Med* 2002;162(22):2614-2620 [FREE Full text] [Medline: [12456234](#)]
16. Anderson KL, Burckhardt CS. Conceptualization and measurement of quality of life as an outcome variable for health care intervention and research. *J Adv Nurs* 1999 Feb;29(2):298-306. [Medline: [10197928](#)]
17. Donker T, Griffiths KM, Cuijpers P, Christensen H. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. *BMC Med* 2009;7:79 [FREE Full text] [doi: [10.1186/1741-7015-7-79](https://doi.org/10.1186/1741-7015-7-79)] [Medline: [20015347](#)]
18. Sirey JA, Bruce ML, Alexopoulos GS, Perlick DA, Raue P, Friedman SJ, et al. Perceived stigma as a predictor of treatment discontinuation in young and older outpatients with depression. *Am J Psychiatry* 2001 Mar;158(3):479-481. [Medline: [11229992](#)]
19. Byrne P. Stigma of mental illness and ways of diminishing it. *Adv Psychiatr Treat* 2000;6:65-72. [doi: [10.1192/apt.6.1.65](https://doi.org/10.1192/apt.6.1.65)]
20. Corrigan P, Watson A. The paradox of self-stigma and mental illness. *Clin Psychol (New York)* 2006;9:35-53. [doi: [10.1093/clipsy.9.1.35](https://doi.org/10.1093/clipsy.9.1.35)]
21. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. *Br J Psychiatry* 2004 Oct;185:342-349 [FREE Full text] [doi: [10.1192/bjp.185.4.342](https://doi.org/10.1192/bjp.185.4.342)] [Medline: [15458995](#)]
22. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). *Aust N Z J Public Health* 2001 Dec;25(6):494-497. [Medline: [11824981](#)]
23. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ* 2004 Jan 31;328(7434):265. [doi: [10.1136/bmj.37945.566632.EE](https://doi.org/10.1136/bmj.37945.566632.EE)] [Medline: [14742346](#)]
24. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. 2nd edition. Geneva: World Health Organization; 2001. URL: http://whqlibdoc.who.int/hq/2001/who_msd_msb_01.6a.pdf [accessed 2012-03-08] [WebCite Cache ID 661KCjMtR]
25. Piccinelli M, Tessari E, Bortolomasi M, Piasere O, Semenzin M, Garzotto N, et al. Efficacy of the alcohol use disorders identification test as a screening tool for hazardous alcohol intake and related disorders in primary care: a validity study. *BMJ* 1997 Feb 8;314(7078):420-424. [Medline: [9040389](#)]
26. Babor T, Campbell R, Room R, Saunders J. *Lexicon of Alcohol and Drug Terms*. Geneva: World Health Organization; 1994.
27. Schmidt S, Mühlan H, Power M. The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. *Eur J Public Health* 2006 Aug;16(4):420-428 [FREE Full text] [doi: [10.1093/eurpub/cki155](https://doi.org/10.1093/eurpub/cki155)] [Medline: [16141303](#)]
28. WHOQOL Group. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Soc Sci Med* 1998 Jun;46(12):1569-1585. [Medline: [9672396](#)]
29. WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *The WHOQOL Group. Psychol Med* 1998 May;28(3):551-558. [Medline: [9626712](#)]

30. Griffiths KM, Christensen H, Jorm AF. Predictors of depression stigma. *BMC Psychiatry* 2008;8:25 [FREE Full text] [doi: [10.1186/1471-244X-8-25](https://doi.org/10.1186/1471-244X-8-25)] [Medline: [18423003](https://pubmed.ncbi.nlm.nih.gov/18423003/)]
31. Kiropoulos LA, Griffiths KM, Blashki G. Effects of a multilingual information website intervention on the levels of depression literacy and depression-related stigma in Greek-born and Italian-born immigrants living in Australia: a randomized controlled trial. *J Med Internet Res* 2011;13(2):e34 [FREE Full text] [doi: [10.2196/jmir.1527](https://doi.org/10.2196/jmir.1527)] [Medline: [21504872](https://pubmed.ncbi.nlm.nih.gov/21504872/)]
32. Hamer RM, Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *Am J Psychiatry* 2009 Jun;166(6):639-641. [doi: [10.1176/appi.ajp.2009.09040458](https://doi.org/10.1176/appi.ajp.2009.09040458)] [Medline: [19487398](https://pubmed.ncbi.nlm.nih.gov/19487398/)]
33. Wright BD, Masters GN. *Rating Scale Analysis*. Chicago, IL: Mesa Press; 1982.
34. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata*. 2nd edition. College Station, TX: Stata Press; 2008.
35. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data* (Springer Series in Statistics). New York, NY: Springer; 2009.
36. Morgenstern J, Longabaugh R. Cognitive-behavioral treatment for alcohol dependence: a review of evidence for its hypothesized mechanisms of action. *Addiction* 2000 Oct;95(10):1475-1490. [Medline: [11070524](https://pubmed.ncbi.nlm.nih.gov/11070524/)]
37. Kiluk BD, Nich C, Babuscio T, Carroll KM. Quality versus quantity: acquisition of coping skills following computerized cognitive-behavioral therapy for substance use disorders. *Addiction* 2010 Dec;105(12):2120-2127. [doi: [10.1111/j.1360-0443.2010.03076.x](https://doi.org/10.1111/j.1360-0443.2010.03076.x)] [Medline: [20854334](https://pubmed.ncbi.nlm.nih.gov/20854334/)]
38. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986 Dec;51(6):1173-1182. [Medline: [3806354](https://pubmed.ncbi.nlm.nih.gov/3806354/)]
39. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009;11(2):e13 [FREE Full text] [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
40. Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 1986 Sep;15(5):451-474. [Medline: [3534875](https://pubmed.ncbi.nlm.nih.gov/3534875/)]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test

CBT: cognitive behavior therapy

WHOQOL: World Health Organization Quality of Life

Edited by G Andersson, A Calear; submitted 07.06.11; peer-reviewed by H Riper, T Berger; comments to author 28.06.11; revised version received 25.09.11; accepted 20.09.11; published 27.06.12.

Please cite as:

Farrer L, Christensen H, Griffiths KM, Mackinnon A

Web-Based Cognitive Behavior Therapy for Depression With and Without Telephone Tracking in a National Helpline: Secondary Outcomes From a Randomized Controlled Trial

J Med Internet Res 2012;14(3):e68

URL: <http://www.jmir.org/2012/3/e68/>

doi: [10.2196/jmir.1859](https://doi.org/10.2196/jmir.1859)

PMID: [22738715](https://pubmed.ncbi.nlm.nih.gov/22738715/)

©Louise Farrer, Helen Christensen, Kathleen M Griffiths, Andrew Mackinnon. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

From Online Randomized Controlled Trials to Participant Preference Studies: Morphing the San Francisco Stop Smoking Site into a Worldwide Smoking Cessation Resource

Ricardo F Muñoz¹, PhD; Adrian Aguilera², PhD; Stephen M Schueller¹, PhD; Yan Leykin³, PhD; Eliseo J Pérez-Stable¹, MD

¹Internet World Health Research Center, Department of Psychiatry, San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, United States

²School of Social Welfare, University of California, Berkeley, Berkeley, CA, United States

³Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:

Ricardo F Muñoz, PhD

Internet World Health Research Center

Department of Psychiatry, San Francisco General Hospital

University of California, San Francisco

1001 Potrero Ave., 7M

San Francisco, CA, 94110

United States

Phone: 1 415 206 5214

Fax: 1 415 206 8942

Email: ricardo.munoz@ucsf.edu

Abstract

Background: Internet interventions have the potential to address many of the health problems that produce the greatest global burden of disease. We present a study illustrating this potential. The Spanish/English San Francisco Stop Smoking Internet site, which yielded quit rates of 20% or more at 12 months in published randomized controlled trials (RCTs), was modified to make it accessible to Spanish- and English-speaking smokers 18 years of age or older anywhere in the world.

Objective: To illustrate that Internet interventions designed to conduct RCTs can be adapted to serve as universal health care resources. We also examine quit rates obtained in the current participant preference study (in which users could choose from all elements tested in previous RCTs) to determine whether they differ from the quit rates found in the RCTs.

Methods: We modified the San Francisco Stop Smoking Internet site so that, instead of being randomly assigned to a specific intervention, participants could personalize the site by choosing among nine site elements (eg, stop smoking guide, reminder emails, journal, mood management intervention, or virtual group). Participants completed a baseline assessment, and reported smoking and mood data at 1-, 3-, 6-, and 12-month follow-ups. We assessed the modified website's reach and outcomes (quit rates), and compared the quit rates of the current participant preference study with those of the previous RCTs.

Results: In the first year of recruitment, 94,158 individuals from 152 countries and territories visited the site; 13,488 participants left some data; 9173 signed consent; 7763 completed the baseline survey; and 1955, 1362, 1106, and 1096 left 1-, 3-, 6-, and 12-month data, respectively. Observed quit rates were 38.1% (n = 668), 44.9% (n = 546), 43.6% (n = 431), and 45.4% (n = 449), respectively. The current participant preference study yielded higher observed quit rates (odds ratio 1.30) than the previous RCT when controlling for individuals' demographic and smoking characteristics.

Conclusions: After strict RCTs are completed, Internet intervention sites can be made into worldwide health intervention resources without reducing their effectiveness.

Trial Registration: Clinicaltrials.gov NCT00721786; <http://clinicaltrials.gov/ct2/show/NCT00721786> (Archived by WebCite at <http://www.webcitation.org/66npiZF4y>)

(*J Med Internet Res* 2012;14(3):e64) doi:[10.2196/jmir.1852](https://doi.org/10.2196/jmir.1852)

KEYWORDS

Internet intervention; smoking cessation; international resources; Spanish; English; outcome study

Introduction

The concept of health care as a basic human right has been advanced by many. Franklin Roosevelt, in his 1944 message to Congress, called for “a second Bill of Rights” that would include “the right to adequate medical care and the opportunity to achieve and enjoy good health” [1]. This goal was included in the United Nations Universal Declaration of Human Rights [2]. However, the concept remains an aspiration that has not been actualized in most of the world. More than 40 years ago, Dr Martin Luther King Jr spoke to the Medical Committee for Human Rights, saying, “Of all the forms of inequality, injustice in health care is the most shocking and inhumane” [3]. We argue that self-help, automated Internet interventions could contribute to making the goal of adequate health care for all a reality. We illustrate this potential by describing how an Internet stop smoking intervention found effective in randomized controlled trials (RCTs) was made available freely to individuals worldwide.

In 2008, the World Health Organization released a report that stated that “Tobacco is the single most preventable cause of death in the world today...Unless urgent action is taken tobacco could kill *one billion people* during this century” [4]. To reduce this figure, smoking cessation interventions must be disseminated widely. There are over one billion smokers in the world today, 80% of whom reside in low- and middle-income countries [5]. Most communities do not offer smoking cessation interventions.

Internet interventions can fill a void in health care by providing evidence-based smoking cessation anytime, anywhere, to anyone with Internet access. Consumable interventions, in which the active agent is spent when administered, are expensive in terms of money, time, or both, so they are rarely used on a large scale. For example, smoking cessation groups consume the counselor’s time: each hour spent can never be used to serve another smoker. Nicotine patches or other pharmacological interventions can be used only once. Thus, each unit of time or money benefits only one individual or group of individuals. This makes such services expensive and thus difficult to establish and sustain. Nonconsumable interventions, such as automated self-help Internet interventions, are a great deal less expensive per user when delivered on a large scale. Though they require a modest amount of funding to host and maintain the website, after the interventions are provided to several thousand individuals (which is very feasible via the Web), the marginal cost (the cost of providing the same services to one more individual) approaches zero. This makes Internet interventions particularly useful if we want to blanket the world with evidence-based health interventions.

The San Francisco Stop Smoking Internet site was developed and tested in RCTs with funding from the California Tobacco-Related Disease Research Program, obtaining quit rates of 20% or more at 12 months [6,7], which are comparable with quit rates of 14%–27% reported for smoking cessation

therapies or smoking cessation groups [7]. Our Internet interventions were extensions of a research program focused on Spanish-speaking smokers [8,9] and were translated and adapted to English to reach a larger number of smokers worldwide. After the grant funding ended for the last RCT, we decided not to close down the site, as would be done with most research grant-funded face-to-face interventions, but rather to keep it active and available as a service to smokers worldwide. We removed the randomization function and most eligibility criteria (except being 18+ years old), and allowed participants to choose any site element(s), thus conducting a participant preference study. These three changes allowed us to assess the dissemination and effectiveness of the intervention on a large scale and to experiment with a working model for how Internet research sites can be made accessible as basic services to any adult in the world with Internet access. This is a clear advantage of Internet interventions over traditional face-to-face outcome studies. When face-to-face trials end, even if they are highly successful, the investigators are almost always unable to continue to provide the same level of service to the public (mainly because staffing costs are prohibitive), leaving them to hope that the knowledge they have gathered will eventually be shared via adoption of their intervention by others. Internet interventions, such as the one described here, can reduce the average length of time that empirically supported interventions take to go from the laboratory to routine use. The US Institute of Medicine reported that it takes an average of 17 years for knowledge from RCTs to be made available to the public [10]. The San Francisco Stop Smoking site was switched from an RCT to an open participant preference site on the day the RCT ended.

This report is the fourth in a line of outcome papers based on the San Francisco Stop Smoking Internet website (www.stopsmoking.ucsf.edu in English; www.dejardefumar.ucsf.edu in Spanish). The first outcome paper [6] reported on the initial set of studies from the website, including single-condition studies and studies comparing two conditions at a time. The second outcome paper [7] compared four conditions with cohort maintenance efforts intended to reduce attrition at follow-up. These efforts included phone calls to the first 500 English-speaking and the first 500 Spanish-speaking participants. The four conditions were as follows: condition 1, The Stop Smoking Guide; condition 2, condition 1 + individually timed email messages; condition 3, condition 2 + a mood management intervention; and condition 4, condition 3 + a virtual group (an asynchronous bulletin board). After these first 1000 participants had been recruited, the site was left active and continued to recruit an additional 16,430 participants, who went through the randomized trial with the same four conditions in a totally automated manner, without the cohort maintenance efforts of the 2009 report [11].

In this paper, we report characteristics of the smokers and quit rates obtained from the sample recruited during the first year of the participant preference site. We compare these data to similar data obtained in the fully automated RCT with 16,430

participants. We hypothesized that providing choice would result in quit rates at least as high as those in the previous RCT.

Methods

Participants

Participants were recruited primarily through a Google Adwords campaign supported by a Google grant awarded to the senior researcher (RFM). The grant allowed us to bid up to US \$1 to place sponsored links (ads) on pages resulting from Google searches relevant to smoking cessation. Individuals using search terms related to smoking were presented with a sponsored link to the English or Spanish version of the website (www.stopsmoking.ucsf.edu or www.dejardefumar.ucsf.edu). Spanish searches tended to have fewer advertisers competing for ad space; therefore, sponsored links for the site were more likely to appear on the first page of Google searches in that language. Others learned about the site through organic searches in Google and other search engines, links from other sites, word of mouth, and media sources.

Study Procedures

The site offered all visitors to the site immediate access to the two major active intervention elements: the Stop Smoking Guide and the Nicotine Replacement Therapy Guide (see below for detailed description), which could be accessed from the home page without the need to sign up. Participants were informed that the site contained a program that is for smokers who are ready to quit; will take approximately 8 weeks to complete; and involves brief questionnaires at 1, 3, 6, and 12 months after enrollment to monitor progress with quitting and usage of the site. (This timing of follow-ups is common in cessation studies; it achieves a good balance between available data and low participant burden). Participants who clicked on a button to enroll in the study were directed to brief demographic and eligibility questionnaires. The sole eligibility criterion was being at least 18 years old. Those eligible and providing an email address received a password to log in to the site. This password also was used to sign the online consent form. Ineligible participants or those not consenting to participate in the research study could access the Stop Smoking Guide (which is itself an evidence-based intervention) and the Nicotine Replacement Therapy Guide from the home page of the website.

On logging in to the website, participants received a baseline questionnaire that assessed demographics, smoking patterns and history, and mood. After completing these assessments, participants were presented with the choice of nine intervention elements. Participants could select any and all of these elements. Unselected elements were suppressed. Participants were then taken to their individualized home page that contained a navigation bar with the elements they had selected. They could access these elements throughout the 12-month study. Participants received emails to complete follow-up assessments at 1, 3, 6, and 12 months after enrolling in the study. These intervals were identical to the ones used in our previous RCTs and thus are directly comparable.

Intervention Elements

The participant preference version of the San Francisco Stop Smoking Internet site allowed users to choose among the following nine elements: (1) the Prequit Checklist, (2) The Stop Smoking Guide (*Guía*), (3) Nicotine Replacement Therapy Guide, (4) Taking Control of Your Life (*Tomando Control*), (5) individually timed email messages, (6) the Mood Management Intervention, (7) the Virtual Group, (8) the Journal, and (9) the Cigarette Counter.

The Prequit Checklist

This 10-item to-do list includes tips on removing smoking-related cues from one's environment (eg, disposing of cigarettes, throwing out smoking supplies, and cleaning items that might smell of tobacco), identifying situations that might lead to relapses, and dealing with those situations.

The Stop Smoking Guide (Guía)

This is a National Cancer Institute evidence-based behavioral intervention originally developed for Spanish-speaking smokers [8,9,12], which provides empirical information about the effects of cigarettes, as well as methods for successful cessation.

Nicotine Replacement Therapy Guide

This guide outlines who should consider nicotine substitutes, information regarding options for nicotine substitutes (such as the nicotine patch, gum, and inhalers), and antidepressant medications used to help people stop smoking.

Taking Control of Your Life (Tomando Control)

This downloadable document helps individuals quit smoking by encouraging them to maintain a healthy mood state by engaging in pleasant, healthy activities [8,12]. The document instructs participants to keep a daily log of the number of positive activities, mood, and the number of cigarettes smoked. These materials were developed for an earlier study [9] and can be found on the National Cancer Institute's site for Research Tested Intervention Programs [13].

Individually Timed Email Messages

These messages are brief tips and encouragement to stop smoking [14], keyed to the participant's self-selected quit date.

The Mood Management Intervention

This 8-lesson cognitive-behavioral mood management course teaches the relationships between thoughts, activities, people, and one's mood; it has been found effective at reducing symptoms of depression and increasing rates of quitting [9,15]. It includes tools to track mood, activities, thoughts, and interpersonal contacts, and to visualize how these are related to cigarettes smoked.

The Virtual Group

This asynchronous bulletin board provides an online forum where participants can post messages and respond to other users' posts.

The Journal

A text box is provided for participants to keep notes on their progress toward their goals. Previous Journal entries can be

viewed via a calendar. They can also be shared with the virtual group.

The Cigarette Counter

Participants indicate the number of cigarettes they smoked the previous day on a visual scale. Results from the cigarette counter are displayed graphically so participants can view their smoking patterns over time.

Measures

A *demographic questionnaire* asked for age, gender, race/ethnicity, education, income, marital status.

A *smoking questionnaire* included age when the person started smoking, age when smoking regularly, number of cigarettes per day, confidence in quitting, smoking exposure, and smoking cessation methods used.

The *Fagerström Test for Nicotine Dependence* (FTND) [16] is a commonly used 6-item test of nicotine dependence, with a range from 0 to 10.

The *Major Depressive Episode (MDE) Screener* (Mood Screener) [17] screens for the presence of the 9 symptoms of current and past MDEs according to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) [18], as well as for criterion C (significant impairment in functioning). This instrument has been shown to have good agreement with the Primary Care Evaluation of Mental Disorders [19,20] and with clinician-administered interviews [21].

The *Center for Epidemiologic Studies Depression* scale (CES-D) [22] is a 20-item self-report scale designed to measure the current level of depressive symptoms.

Self-reported 7-day point prevalence abstinence was defined as a “no” response to “Have you smoked 1 or more cigarettes in the last 7 days?” at 1-, 3-, 6-, and 12-month follow-ups. This is the main measure used in the previously published RCT [7].

Statistical Analyses

We constructed two logistic regression models to predict quit rates. Predictors were age, sex, education, language, FTND score, number of cigarettes smoked per day, age when the person started smoking, number of quit attempts, use of quit aids,

CES-D score, and presence of lifetime MDEs. Variables were entered stepwise in groups: demographic variables, smoking variables, and depression variables. One model assessed differences between all participants in the current participant preference study and the previous RCT. Another model tested differences only among individuals who would have met the more-stringent requirements of the previous RCT (smoked more than 5 cigarettes a day, intending to quit in the next month, used email at least once weekly, and was 18+ years of age). Chi-square and independent sample *t* tests were also conducted with descriptive data to make comparisons as in the case of attrition data, as well as to compare between the Spanish- and the English-speaking subsamples.

Because the website was open to anyone 18 years of age or older, some participants were nonsmokers at baseline. We analyzed outcome excluding these nonsmokers.

Results

Sample

As seen in [Figure 1](#), during a 1-year period (September 19, 2008 through September 18, 2009), 94,158 ($n = 39,267$ English-speaking, $n = 54,891$ Spanish-speaking) individuals from 152 countries and territories visited the site; 13,488 ($n = 3372$ English-speaking, $n = 10,116$ Spanish-speaking) participants left data; 9173 ($n = 2352$ English-speaking, $n = 6821$ Spanish-speaking) signed consent; and 7763 ($n = 1825$ English-speaking; $n = 5938$ Spanish-speaking) participants completed a baseline survey. Participants completing baseline represented 124 countries and territories. English speakers came from 111 countries, of which the top 3 were India ($n = 442$), the United States ($n = 240$), and South Africa ($n = 182$). Spanish speakers came from 51 countries, of which the top 3 were Spain ($n = 2573$), Argentina ($n = 1114$), and Mexico ($n = 547$). [Table 1](#) further describes sample characteristics. [Table 2](#) describes the smoking and depression characteristics of participants who identified as smokers or reported smoking 1+ cigarettes at baseline ($n = 7353$). As can be seen from [Table 2](#), Spanish-language participants appeared to be less addicted to nicotine; they were also far less likely to report using quit aids in the past 6 months and more likely to screen positive for current MDE.

Table 1. Demographics of consenting participants who completed baseline assessment.

Characteristic	Smokers			Nonsmokers
	All (n = 7353)	English-speaking (n = 1693)	Spanish-speaking (n = 5660)	All (n = 410)
Sex ^a	n = 7314	n = 1688	n = 5626	n = 410
Men	4024 (55.02%)	1099 (65.11%)	2925 (51.99%)	208 (50.7%)
Women	3290 (44.98%)	589 (34.9%)	2701 (48.01%)	202 (49.3%)
Ethnicity ^{a,b}	n = 7286	n = 1658	n = 5628	n = 406
Hispanic/Latino	5538 (76.01%)	185 (11.2%)	5232 (92.96%)	273 (67.2%)
Not Hispanic/Latino	1752 (23.99%)	1356 (81.79%)	396 (7.04%)	133 (32.8%)
Race ^{a,c}	n = 7097	n = 1478	n = 5619	n = 407
European descent	4435 (62.49%)	617 (41.7%)	3818 (67.95%)	280 (70.3%)
Asian descent	531 (7.5%)	520 (35.2%)	11 (0.2%)	19 (5%)
African descent	48 (1%)	42 (3%)	6 (0.1%)	2 (1%)
Indigenous descent	20 (0.3%)	3 (0.2%)	17 (0.3%)	0 (0%)
Other/multiethnic	1292 (18.20%)	253 (17.1%)	1019 (18.13%)	76 (18.6%)
Mestizo	771 (10.9%)	23 (2%)	748 (13.3%)	23 (5.7%)
Education ^{a,b}	n = 7216	n = 1664	n = 5552	n = 406
High school or less	1688 (23.39%)	296 (17.8%)	1392 (25.07%)	62 (15%)
Some college	2626 (36.39%)	428 (25.7%)	2198 (39.59%)	129 (31.8%)
College graduate	2126 (29.46%)	622 (37.4%)	1504 (27.09%)	139 (34.2%)
Graduate degree	776 (10.75%)	318 (19.1%)	458 (8.2%)	76 (18%)
Marital status ^{a,d}	n = 7345	n = 1690	n = 5655	n = 410
Single	2647 (36.04%)	722 (42.7%)	1925 (34.04%)	114 (28.0%)
Cohabiting	1224 (16.66%)	195 (11.5%)	1029 (18.20%)	83 (20.2%)
Married	2674 (36.41%)	626 (37.0%)	2046 (36.18%)	167 (40.7%)
Separated	354 (4.8%)	55 (3%)	299 (5.3%)	16 (4%)
Divorced	395 (5.4%)	82 (5%)	313 (5.5%)	25 (6%)
Widowed	53 (1%)	10 (1%)	43 (1%)	4 (1%)

^a Difference between English- and Spanish-speaking participants is significant at $P < .001$ level.^b Difference between smokers and nonsmokers is significant at $P < .001$ level.^c Difference between smokers and nonsmokers is significant at $P = .003$ level.^d Difference between smokers and nonsmokers is significant at $P = .03$ level.

Figure 1. Progression of participants through the San Francisco Stop Smoking participant preference site, with follow-up at 1 month (1m), 3 months (3m), 6 months (6m), and 12 months (12m).

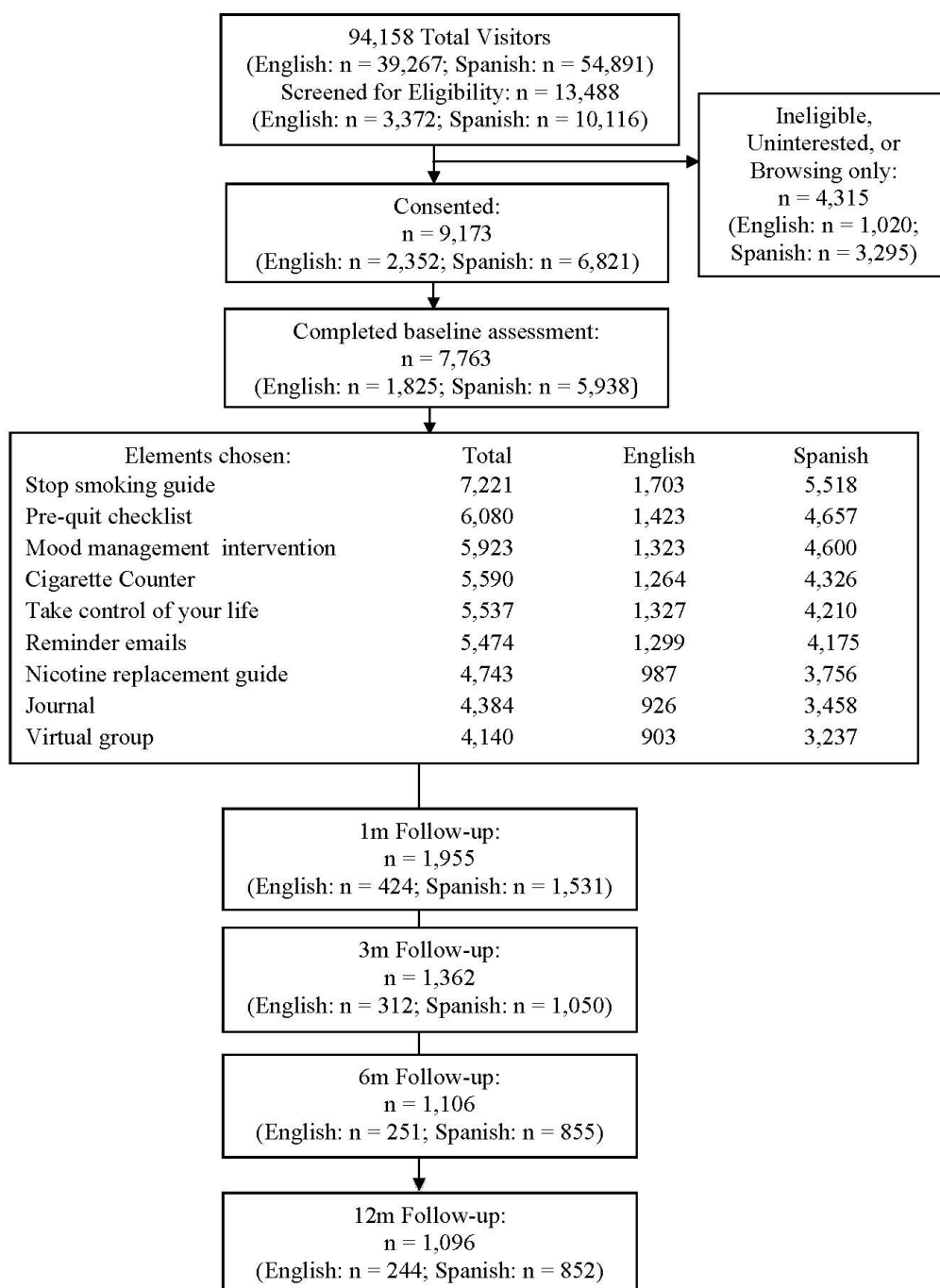


Table 2. Smoking and clinical characteristics of consenting smokers.

Characteristic	All (n = 7353)	English-speaking (n = 1693)	Spanish-speaking (n = 5660)	P value (English vs Spanish)
Smoking history, mean (SD)				
Age (years)	35.0 (10.7)	33.6 (10.7)	35.8 (10.7)	<.001
Age (years), first cigarette	15.8 (3.3)	16.6 (4.1)	15.5 (3.0)	<.001
Age(years) regular smoker	18.8 (4.4)	19.0 (4.3)	18.8 (4.4)	.07
Cigarettes per day	18.1 (10.0)	16.7 (10.0)	18.5 (10.0)	<.001
FTND ^a score	6.5 (2.6)	7.4 (2.0)	6.2 (2.7)	<.001
Quit confidence	6.6 (2.3)	6.7 (2.3)	6.6 (2.3)	.04
Methods used to quit in last 6 months, n (%)				
Nicotine gum	884 (12.0%)	235 (13.9%)	649 (11.5%)	0.008
Nicotine patch	580 (7.9%)	220 (13.0%)	360 (6.4%)	<.001
Nicotine inhaler	60 (1%)	41 (2%)	19 (0.3%)	<.001
Nicotine spray	18 (0.2%)	10 (1%)	8 (0.1%)	.003
Bupropion	223 (3.0%)	64 (4%)	159 (2.8%)	.04
Varenicline	170 (2.3%)	37 (2%)	133 (2.3%)	.78
Other antidepressant	103 (1.4%)	29 (2%)	74 (1%)	.24
Stop smoking group	124 (1.7%)	27 (2%)	97 (2%)	.83
Hypnosis	116 (1.6%)	56 (3%)	60 (1%)	<.001
Acupuncture	126 (1.7%)	29 (2%)	97 (2%)	1.00
Motivational tapes	126 (1.7%)	57 (3%)	69 (1%)	<.001
Other self-help	321 (4.4%)	118 (7.0%)	203 (3.6%)	<.001
Other websites	188 (2.6%)	51 (3%)	137 (2.4%)	.19
Prayer	271 (3.7%)	119 (7.0%)	152 (2.7%)	<.001
Consultation with doctor	239 (3.3%)	65 (4%)	174 (3.1%)	.12
Other	328 (4.5%)	108 (6.4%)	220 (3.9%)	<.001
None	5047 (68.64%)	997 (58.9%)	4050 (71.55%)	<.001
Depression variable, n (%)				
Major depressive episodes				
No history	4968 (67.56%)	1143 (67.51%)	3825 (67.58%)	.80
Past only	1027 (13.97%)	283 (16.7%)	744 (13.1%)	<.001
Current	1358 (18.47%)	267 (15.8%)	1091 (19.3%)	.002
CES-D ^b score, mean (SD)	17.7 (12.2)	17.9 (11.9)	17.7 (12.3)	.35

^a Fagerström Test for Nicotine Dependence.^b Center for Epidemiologic Studies Depression scale.

Table 3. Self-reported 7-day smoking abstinence, for all smokers and, separately, for smokers eligible for the previous RCT^a (% quit).

	All (n = 7353)	English-speaking (n = 1693)	Spanish-speaking (n = 5660)	P value (English vs Spanish)
Observed quit rate				
All smokers				
Month 1	668 (38.1%)	141 (38.8%)	527 (37.9%)	.77
Month 3	546 (44.9%)	115 (42.9%)	431 (45.5%)	.49
Month 6	431 (43.6%)	100 (47.8%)	331 (42.5%)	.19
Month 12	449 (45.4%)	104 (50.0%)	345 (44.2%)	.14
RCT eligible				
Month 1	587 (40.4%)	127 (40.1%)	460 (40.5%)	.90
Month 3	466 (45.4%)	101 (43.0%)	365 (46.1%)	.42
Month 6	378 (45.9%)	87 (49%)	291 (45.0%)	.35
Month 12	377 (45.8%)	93 (51%)	284 (44.2%)	.11
Missing = smoking^b				
All smokers				
Month 1	668 (9.1%)	141 (8.3%)	527 (9.3%)	.23
Month 3	546 (7.4%)	115 (6.8%)	431 (7.6%)	.27
Month 6	431 (5.9%)	100 (5.9%)	331 (5.8%)	.96
Month 12	449 (6.1%)	104 (6.1%)	345 (6.1%)	.96
RCT eligible				
Month 1	587 (9.8%)	127 (8.8%)	460 (10.1%)	.17
Month 3	466 (7.8%)	101 (7.0%)	365 (8.0%)	.24
Month 6	378 (6.3%)	87 (6%)	291 (6.4%)	.71
Month 12	377 (6.3%)	93 (7%)	284 (6.2%)	.76

^a Smokers who met eligibility criteria set for the 2009 randomized controlled trial (RCT) [7,11].

^b Assuming that every participant not reporting data is still smoking.

Attrition

Of the 7763 consenting participants who completed the baseline survey, 1955 (25.18%), 1362 (17.54%), 1106 (14.25%), and 1096 (14.12%) left 1-, 3-, 6-, and 12-month data, respectively. Most participants (4914, 63.30%) did not return to the site for any of the follow-ups. Of the 2855 (36.78%) who left follow-up data, 1423 (49.84%) responded to one, 629 (22.0%) responded to two, 367 (12.8%) responded to three, and 436 (15.3%) responded to all four follow-ups. These rates were numerically lower than those obtained in the RCT. We found no differences in the proportion of follow-ups completed by English or Spanish speakers ($\chi^2_4 = 6.38$, $P = .17$). There was a significant difference in the proportion of completed follow-ups between women and men ($\chi^2_4 = 37.88$, $P < .001$), with the largest difference between those who completed no follow-ups (2798, 66.12% men vs 2105, 60.28% women).

Quit Rates (7-Day Abstinence)

In Table 3, we report both the observed quit rates (based on those reporting data at each follow-up) and the missing = smoking quit rates (that is, assuming that every participant not

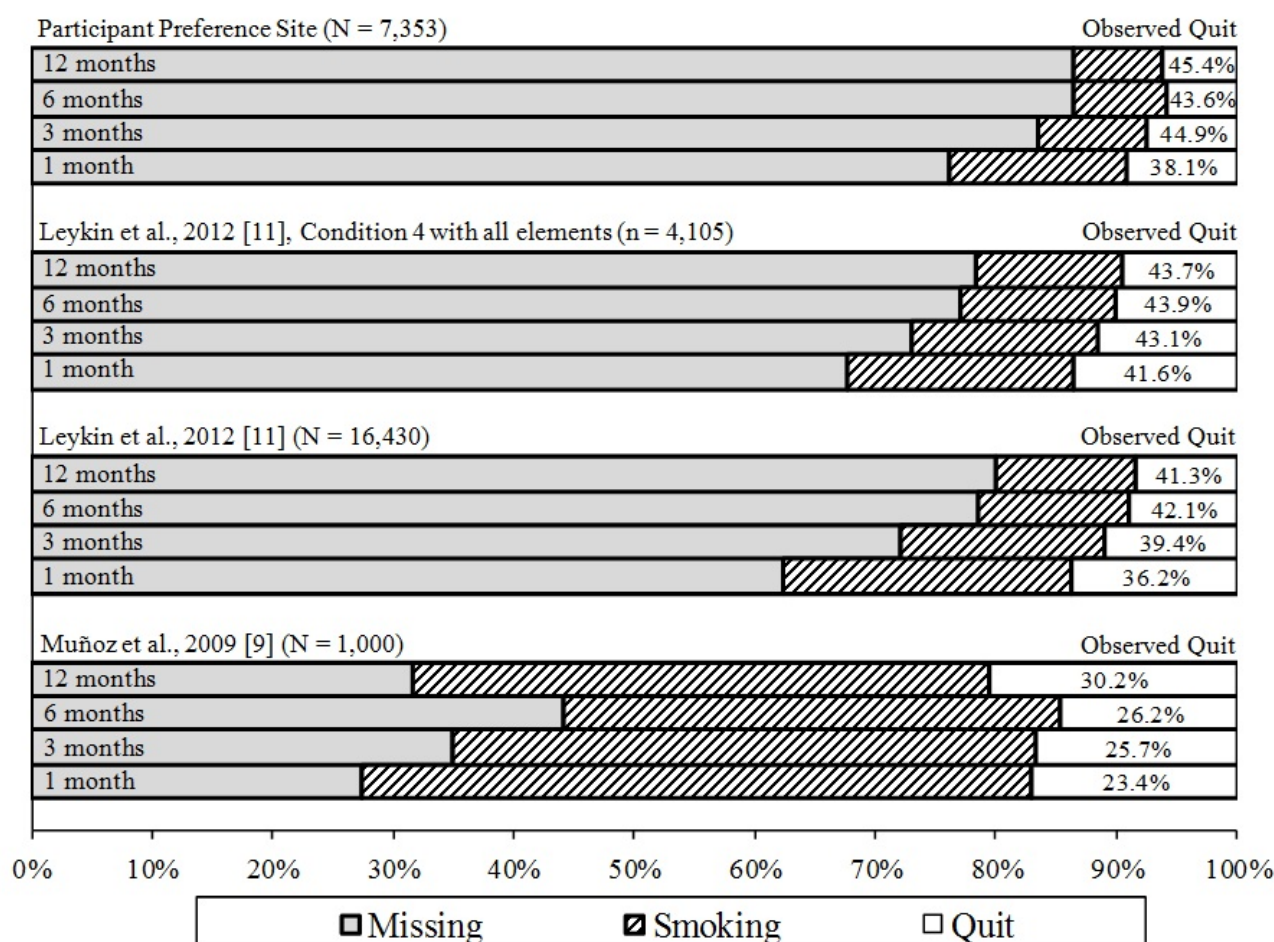
reporting data is still smoking) for participants in the participant preference trial. We ran a logistic regression model to compare the observed quit rates between the previous RCT and the current participant preference study controlling for demographic, smoking, and depression variables. We found that participants in the current participant preference study reported higher quit rates than participants in the previous RCT (Wald $\chi^2_1 = 10.43$, $P < .001$, odds ratio [OR] 1.30). With both samples combined, those with more education were more likely to quit (Wald $\chi^2_1 = 14.51$, $P < .001$, OR 1.14); English-speaking participants were less likely to do so (Wald $\chi^2_1 = 6.52$, $P = .01$, OR 0.81). The only smoking-related variable related to quit rates was nicotine dependence, such that higher dependence corresponded to lower quit rates (Wald $\chi^2_1 = 13.30$, $P < .001$, OR 0.94). Both levels of depressive symptoms (Wald $\chi^2_1 = 20.80$, $P < .001$, OR 0.99) and history of MDEs (Wald $\chi^2_1 = 8.69$, $P = .003$, OR 0.79) predicted lower quit rates in participants. The largest change in R² resulted from the addition of the depression variables.

Figure 2 displays the quit rates obtained from this study along with those obtained in previous RCTs. The observed 1-, 3-, 6-,

and 12-month quit rates among smokers in the current study were 38.1%, 44.9%, 43.6%, and 45.4%, respectively. Restricting the sample only to participants who met eligibility criteria for the most recent, fully automated RCT [11] produced observed quit rates of 40.4% (n=587), 45.4% (n=378), 45.9% (n=378), and 45.8% (n=377) respectively. In that RCT, quit rates at the same time points for participants in condition 4, who received all elements (the most reasonable comparison), were 41.6% (552/1328), 43.1% (475/1103), 43.9% (414/943), and 43.7%

(386/884). Note that in the 2009 RCT, we instituted cohort maintenance methods to obtain as complete a set of follow-up data as possible by phoning participants who did not respond to automated emails. This resulted in follow-up rates as high as 69% at 12 months. The fully automated RCT reported in the Leykin et al [11] and the participant preference site results did not include these phone calls. Thus, the observed quit rates are based on a much smaller proportion of respondents. This needs to be taken into account when interpreting these figures.

Figure 2. Observed quit rates by study for the San Francisco Stop Smoking participant preference study and previous randomized controlled trials (RCTs). The outcomes of the current participant preference site, as compared with the RCT with live cohort maintenance follow-up methods ([7], N = 1000), a fully automated RCT without live follow-up ([11], N = 16,430), and condition 4, the condition in that same automated RCT that included all site elements, which approximates the current site, but without the participant preference component ([11], n = 4105). Shaded areas indicate missing data. Observed Quit is defined as the number reporting not smoking divided by number reporting smoking status (smoking or not smoking). Thus, it does not impute values for missing data.



Discussion

This study illustrates that it is feasible to blanket the world with self-help automated Internet interventions for major global health problems. We showed that turning an efficacious Internet-based RCT into a participant preference study site immediately after the RCT ended resulted in quit rates similar to or better than those found in the RCT. Opening up an Internet intervention research site after completion of an Internet outcome study allows people worldwide to continue to use it,

and the outcomes can be similar to those obtained within a strict RCT context.

We found higher observed quit rates among those in the participant preference study. In addition, several variables predicted higher levels of quit rates. Specifically, higher levels of education, visiting the Spanish version of the site, lower levels of nicotine dependence, lower depressive symptoms, and no history of MDEs were all more highly associated with increased quit rates. We found the higher quit rates related to the Spanish site intriguing. Perhaps because Spanish-speaking smokers were less likely to have used other smoking cessation

resources (see Table 2), the Internet site was more likely to be of help to them in part by helping those who had an easier time quitting. English speakers able to quit more easily may have already quit using other methods. This hypothesis is supported by the fact that English speakers had significantly higher FTND nicotine dependence scores (see Table 2). If this is correct, it is one more argument for blanketing the world with evidence-based Internet interventions, so that individuals in locations that have no health resources can have access to at least a basic level of care, which can be effective in addressing the targeted health problem. Note, in Table 2, that 2 out of 3 smokers had used nothing prior to our Internet intervention. Only 3.3% had consulted with a doctor, and 1.7% had attended a smoking cessation group. The most used aid was nicotine gum, but even that method was used by only 12% of smokers.

These findings illustrate how an automated stop smoking site can be used to provide a basic level of health care. Just as visits to professionals might help refer participants to additional or more-specialized services, an Internet stop smoking site can serve as a basic level of health care that could link participants to additional services if the basic level is not enough to deal with their health issues.

To help the reader consider the comparative advantage of this Internet intervention, we estimated the contribution of the site relative to comparable spending on alternative interventions. The San Francisco Stop Smoking site provided free access to the Stop Smoking Guide and Nicotine Replacement Therapy Guide to 94,158 visitors to the site. This would be comparable with providing that many printed brochures and mailing them to smokers all over the world. As with brochures, of course, there is no guarantee they would be read and acted upon. To serve the 9173 smokers who signed consent with other widely used interventions, it would have cost US \$2,008,520 in nicotine patches (at \$3.91 per day at 8 weeks per person), \$2,984,527 in nicotine gum (at \$5.81 per day and 8 weeks per person), \$2,208,858 in bupropion (at \$4.30 per day and 8 weeks per person) [23], and \$1,174,144 in group counseling (at \$16 per hour of counselor time once per week [24]. In comparison, the cost of maintaining the self-help, automated site for 1 year was approximately US \$50,000 (hosting and maintenance costs, and the cost of a half-time bilingual staff person to handle email inquiries and check the site periodically).

To have the greatest impact in reducing unnecessary suffering due to health problems, we recommend that international agencies and Internet intervention researchers prioritize (1) the

health problems producing the greatest burden of disease globally, such as smoking, which is the number one cause of preventable death, and depression, which is the number one cause of disability worldwide [25], (2) the health problems that can be addressed using behavior change interventions, and (3) the languages that reach the largest number of people via the Web (currently English, Chinese, and Spanish). At the same time, we want to avoid duplicating the problem of orphan drugs, which are neglected because they work for less-common disorders. To do so, we encourage researchers with the expertise to address less-prevalent health issues and languages spoken by smaller populations to focus on these issues as well. For health care to be a right for all, it should include even numerically small minorities.

If we blanketed the world with automated, self-help Internet interventions such as the San Francisco Stop Smoking Internet site as a basic level of health care, many people without appropriate health care access could benefit from the online intervention. Communities with more resources would be able to add more-expensive and consumable services, in this case, additional smoking cessation aids. For example, they could provide smoking cessation groups that use the website as an adjunct, or they could provide nicotine patches, gum, and other pharmacological aids to smokers using the website. These communities might also learn how to better link basic, scalable levels of health care, such as Internet tools, to more-involved forms of care. The stop smoking site might have served as a portal for identifying those interested and motivated to quit, providing basic skills that apparently benefitted many, and providing motivation for further care for those who need it. Linked resources could help alleviate concerns that those who visit Internet sites and do not respond will feel discouraged from seeking out future support, especially if they are explicitly told that this is something that will help many, but not all, and that some will need more-individualized treatment.

We have presented a successful example of developing an Internet health intervention, testing it in strict RCTs, and immediately disseminating it to a large worldwide population. We suggest that international health institutions, such as the World Health Organization, establish official, adequately funded programs to systematically develop, test, and disseminate evidence-based, self-help, automated, and multilingual interventions to all who need them. This will contribute to the goal of making health care a freely accessible right for people worldwide.

Acknowledgments

This research program was supported by the Tobacco-Related Disease Research Program, the University of California, Office of the President, the Google Adwords Grants Program, the Brin Wojcicki Foundation, the National Cancer Institute for Redes en Acción (U01CA86117) and by National Institute of Mental Health grants 5T32MH018261-27 and 5K08MH091501. We would also like to thank the Center for Health and Community at University of California, San Francisco, the Information Services Unit, Leslie Lenert, Carlos Penilla, Eduardo Ballesteros, Thomas Manley, and Veronica Pitbladdo for their contributions to this line of research. All assessment surveys were conducted using the Web survey system Illume by DatStat, Inc.

Conflicts of Interest

None declared.

References

1. Moss WG. George Mason University's History News Network. 2009 Nov 07. Is Health Care a Human Right? URL: <http://hnn.us/articles/119553.html> [accessed 2011-05-24] [WebCite Cache ID 5yuZLH6gu]
2. United Nations. 1948. The Universal Declaration of Human Rights URL: <http://www.un.org/en/documents/udhr/> [accessed 2011-08-03] [WebCite Cache ID 60fgfSLfs]
3. Lewis LT. Huffington Post. 2011 Jan 20. Honoring a Dream for Health Care Justice URL: http://www.huffingtonpost.com/dr-l-toni-lewis/honoring-a-dream-for-heal_b_811600.html [accessed 2011-05-24] [WebCite Cache ID 5yuZYmwXH]
4. World Health Organization. 2008. WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER Package URL: http://www.who.int/tobacco/mpower/mpower_report_full_2008.pdf [accessed 2011-08-03] [WebCite Cache ID 60fdnd1Xa]
5. WHO Media Centre. World Health Organization. 2011 Feb. Tobacco Fact Sheet no 339 URL: <http://www.who.int/mediacentre/factsheets/fs339/en/index.html> [accessed 2011-05-24] [WebCite Cache ID 5yuZhNW2E]
6. Muñoz RF, Lenert LL, Delucchi K, Stoddard J, Perez JE, Penilla C, et al. Toward evidence-based Internet interventions: A Spanish/English Web site for international smoking cessation trials. *Nicotine Tob Res* 2006 Feb;8(1):77-87. [doi: [10.1080/14622200500431940](https://doi.org/10.1080/14622200500431940)] [Medline: [16497602](https://pubmed.ncbi.nlm.nih.gov/16497602/)]
7. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. *Nicotine Tob Res* 2009 Sep;11(9):1025-1034. [doi: [10.1093/ntr/ntp090](https://doi.org/10.1093/ntr/ntp090)] [Medline: [19640833](https://pubmed.ncbi.nlm.nih.gov/19640833/)]
8. Pérez-Stable EJ, Sabogal F, Marín G, Marín BV, Otero-Sabogal R. Evaluation of "Guía para Dejar de Fumar," a self-help guide in Spanish to quit smoking. *Public Health Rep* 1991;106(5):564-570. [Medline: [1910191](https://pubmed.ncbi.nlm.nih.gov/1910191/)]
9. Muñoz RF, Marín BV, Posner SF, Pérez-Stable EJ. Mood management mail intervention increases abstinence rates for Spanish-speaking Latino smokers. *Am J Community Psychol* 1997 Jun;25(3):325-343. [Medline: [9332966](https://pubmed.ncbi.nlm.nih.gov/9332966/)]
10. Institute of Medicine. Crossing the Quality Chasm: A New Health Care System for the 21st Century. Washington, DC: National Academy Press; 2001.
11. Leykin Y, Aguilera A, Torres LD, Pérez-Stable EJ, Muñoz RF. Interpreting the outcomes of automated internet-based randomized trials: example of an International Smoking Cessation Study. *J Med Internet Res* 2012;14(1):e5 [FREE Full text] [doi: [10.2196/jmir.1829](https://doi.org/10.2196/jmir.1829)] [Medline: [22314016](https://pubmed.ncbi.nlm.nih.gov/22314016/)]
12. National Cancer Institute. Guía para Dejar de Fumar. NIH Publication No. 02-3001. Washington, DC: US Government Printing Office; 2002.
13. National Cancer Institute. 2010 Mar 7. Programa Latino para Dejar de Fumar URL: <http://rtips.cancer.gov/rtips/programDetails.do?programId=105455> [accessed 2011-12-03] [WebCite Cache ID 63f6vEOU6]
14. Lenert L, Muñoz RF, Perez JE, Bansod A. Automated e-mail messaging as a tool for improving quit rates in an internet smoking cessation intervention. *J Am Med Inform Assoc* 2004;11(4):235-240 [FREE Full text] [doi: [10.1197/jamia.M1464](https://doi.org/10.1197/jamia.M1464)] [Medline: [15064291](https://pubmed.ncbi.nlm.nih.gov/15064291/)]
15. Hall SM, Muñoz RF, Reus VI, Sees KL, Duncan C, Humfleet GL, et al. Mood management and nicotine gum in smoking treatment: a therapeutic contact and placebo-controlled study. *J Consult Clin Psychol* 1996 Oct;64(5):1003-1009. [Medline: [8916629](https://pubmed.ncbi.nlm.nih.gov/8916629/)]
16. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br J Addict* 1991 Sep;86(9):1119-1127. [Medline: [1932883](https://pubmed.ncbi.nlm.nih.gov/1932883/)]
17. Muñoz RF. University of California, San Francisco, School of Medicine, Latino Mental Health Research Program. The Mood Screener: a Major Depression Episode Screener URL: <http://www.medschool.ucsf.edu/latino/manuals.aspx> [accessed 2012-04-17] [WebCite Cache ID 670AnBt4y]
18. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR. 4th edition, text revision. Washington, DC: American Psychiatric Association; 2000.
19. Muñoz RF, McQuaid JR, González GM, Dimas J, Rosales VA. Depression screening in a women's clinic: using automated Spanish- and English-language voice recognition. *J Consult Clin Psychol* 1999 Aug;67(4):502-510. [Medline: [10450620](https://pubmed.ncbi.nlm.nih.gov/10450620/)]
20. Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA* 1994 Dec 14;272(22):1749-1756. [Medline: [7966923](https://pubmed.ncbi.nlm.nih.gov/7966923/)]
21. Vázquez FL, Muñoz RF, Blanco V, López M. Validation of Muñoz's Mood Screener in a nonclinical Spanish population. *Eur J Psychol Assess* 2008;24:57-64. [doi: [10.1027/1015-5759.24.1.57](https://doi.org/10.1027/1015-5759.24.1.57)]
22. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1(3):385. [doi: [10.1177/01462167700100306](https://doi.org/10.1177/01462167700100306)]
23. Schroeder SA. What to do with a patient who smokes. *JAMA* 2005 Jul 27;294(4):482-487. [doi: [10.1001/jama.294.4.482](https://doi.org/10.1001/jama.294.4.482)] [Medline: [16046655](https://pubmed.ncbi.nlm.nih.gov/16046655/)]
24. Hall SM, Lightwood JM, Humfleet GL, Bostrom A, Reus VI, Muñoz R. Cost-effectiveness of bupropion, nortriptyline, and psychological intervention in smoking cessation. *J Behav Health Serv Res* 2005;32(4):381-392. [Medline: [16215448](https://pubmed.ncbi.nlm.nih.gov/16215448/)]

25. Murray CJL, Lopez AD. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability From Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020. Cambridge, MA: Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996.

Abbreviations

CES-D: Center for Epidemiologic Studies Depression scale

FTND: Fagerström Test for Nicotine Dependence

MDE: major depressive episode

OR: odds ratio

RCT: randomized controlled trial

Edited by R Tait, H Christensen, A Callear; submitted 26.05.11; peer-reviewed by A Hatchett, P Anderson; comments to author 07.07.11; revised version received 22.08.11; accepted 27.10.11; published 27.06.12.

Please cite as:

Muñoz RF, Aguilera A, Schueller SM, Leykin Y, Pérez-Stable EJ

From Online Randomized Controlled Trials to Participant Preference Studies: Morphing the San Francisco Stop Smoking Site into a Worldwide Smoking Cessation Resource

J Med Internet Res 2012;14(3):e64

URL: <http://www.jmir.org/2012/3/e64/>

doi: [10.2196/jmir.1852](https://doi.org/10.2196/jmir.1852)

PMID: [22739225](https://pubmed.ncbi.nlm.nih.gov/22739225/)

©Ricardo F Muñoz, Adrian Aguilera, Stephen M Schueller, Yan Leykin, Eliseo J Pérez-Stable. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Working Alliance in Online Cognitive Behavior Therapy for Anxiety Disorders in Youth: Comparison With Clinic Delivery and its Role in Predicting Outcome

Renee E E Anderson¹, PhD; Susan H Spence², PhD; Caroline L Donovan², PhD; Sonja March³, PhD; Samantha Prosser², PhD; Justin Kenardy³, PhD

¹School of Psychology, University of Queensland, Brisbane, QLD, Australia

²Griffith Health Research Institute, Griffith University, Nathan, Australia

³Centre of National Research on Disability and Rehabilitation, School of Medicine, University of Queensland, Brisbane, QLD, Australia

Corresponding Author:

Susan H Spence, PhD
Griffith Health Research Institute
Griffith University
170 Kessels Rd
Nathan, QLD 4111
Australia
Phone: 61 737355447
Fax: 61 737357507
Email: s.spence@griffith.edu.au

Abstract

Background: Substantial evidence exists that positive therapy outcomes are related to the therapist–client working alliance.

Objectives: To report two studies that examined (1) the quality of the working alliance in online cognitive behavior therapy (CBT), with minimal therapist contact, for anxiety disorders in youth, and (2) the role of working alliance and compliance in predicting treatment outcome.

Methods: Study 1 participants were 73 adolescents aged 12 to 18 years who met diagnostic criteria for an anxiety disorder, plus one or more of their parents. Participants were randomly assigned to clinic or online delivery of CBT, with working alliance being assessed for youth and parents after session 3. Study 2 participants were 132 children and adolescents aged 7 to 18 years who met diagnostic criteria for an anxiety disorder, plus one or more of their parents. Youths and parents participated in a minimally therapist-assisted online CBT program supported by brief, weekly emails and a single, short phone call.

Results: Study 1 revealed a strong working alliance for both online and clinic CBT, with no significant difference in working alliance between conditions for adolescents ($F_{1,73} = 0.44$, $P = .51$, $\eta_p^2 = 0.006$, Cohen $d = 0.15$). Parents also reported high working alliance in both conditions, although a slight but significantly higher working alliance in clinic-based therapy ($F_{1,70} = 6.76$, $P = .01$, $\eta_p^2 = 0.09$, Cohen $d = 0.64$). Study 2 showed a significant and substantial decrease in anxiety symptoms following online therapy ($P < .001$ for all outcome measures). Adolescents improved significantly more in overall functioning when working alliance (beta = .22, $t_{79} = 2.21$, $P = .03$) and therapy compliance (beta = .22, $t_{84} = 2.22$, $P = .03$) were higher, with working alliance also predicting compliance (beta = .38, $F_{1,80} = 13.10$, $P = .01$). No such relationships were evident among younger children.

Conclusions: Working alliance is important in determining clinical outcome for online treatment for anxiety among adolescents, with minimal therapist assistance, although this was not the case for younger children.

Trial Registration: Australian New Zealand Clinical Trials Registry: ACTRN12611000900910; http://www.anzctr.org.au/trial_view.aspx?ID=343375 (Archived by WebCite at <http://www.webcitation.org/674C4N3JJ>)

(*J Med Internet Res* 2012;14(3):e88) doi:[10.2196/jmir.1848](https://doi.org/10.2196/jmir.1848)

KEYWORDS

Anxiety; online therapy; children; adolescents; working alliance

Introduction

Anxiety disorders among children and adolescents are prevalent and can result in significant deleterious consequences. Cognitive behavior therapy (CBT) has been shown to be highly efficacious in treating youth anxiety within a clinic-based setting [1,2] and, although only a handful of studies have examined the efficacy of an Internet-based treatment approach to youth anxiety, the results are encouraging and demonstrate similar remission rates to clinic-based CBT [3-5]. Yet, despite the impressive efficacy of CBT and regardless of the mode of delivery, not all young people respond positively to CBT for anxiety disorders. Indeed, approximately 43.5% of children continue to meet diagnostic criteria for a primary anxiety disorder following clinic-based CBT for an anxiety disorder [6], with similar rates evident in the few studies that have used either full or partial Internet delivery of treatment [3,4,7]. Exploring factors that might influence therapy outcome is therefore important, so that existing treatments can be improved upon and more effective treatment approaches can be developed. One factor that has been widely investigated with respect to its relationship to treatment outcome, at least within the adult, clinic-based treatment literature, is that of working alliance.

Working alliance, also referred to as therapeutic alliance, relates to the quality and nature of the therapist–client interaction, the collaborative approach toward the tasks and goals of treatment, and the personal bond or attachment that emerges in therapy [8,9]. There is substantial evidence that a strong working alliance between therapist and client is associated with positive psychotherapy outcomes [10,11]. A positive working alliance is proposed to result in increased motivation for and greater perseverance with task completion, reflecting an increased sense of optimism for a positive outcome and desire to please the therapist. In turn, greater completion of therapy tasks is proposed to result in greater acquisition of knowledge and skills, and a stronger sense of mastery [12]. In addition, a strong working alliance in CBT is proposed to facilitate positive cognitive change, provide disconfirming evidence of dysfunctional cognitions, and reinforce functional interpersonal behavior [13]. A recent meta-analytic review of 38 studies examining the association between working alliance and outcome in face-to-face child psychotherapy suggested a small but significant relationship with an average effect size of 0.14 [14]. Interestingly, and of relevance to the present studies, the effect size was smaller for adolescents (0.10) than for children (0.20), and also smaller for internalizing (0.10) than for externalizing disorders (0.22).

The emergence of online and computer-based delivery of psychological treatments has raised issues of whether a strong working alliance can be established in the absence of face-to-face contact with a therapist, and whether the quality of the online working alliance predicts treatment outcomes in this form of therapy. Indeed, online approaches have been criticized for their reduced capacity to establish a strong working alliance given the absence of nonverbal interpersonal cues, difficulty in monitoring client understanding of concepts, and limited ability to provide timely corrective feedback [15]. Despite the reservations expressed by some authors, however, empirical

studies with adults have shown that a strong working alliance can indeed be established in online therapy and is comparable in strength with that of clinic-based treatment [16,17].

Regarding working alliance and outcome for online CBT for youth anxiety, only one study could be identified, and this was limited to a condition in which therapy was delivered primarily within the clinic but supplemented by computer-assisted technology [3]. The authors reported no significant differences on youth therapeutic alliance scores when therapy was partially computer delivered compared with a full clinic therapy format. It remains to be determined whether a strong working alliance can be produced when CBT for youth anxiety is delivered fully online with minimal therapist assistance and whether working alliance in this context predicts therapy outcome. Furthermore, given that treatment for anxiety disorders frequently involves parents, the quality of the working alliance must be considered from the perspective of both parties. The limited evidence available suggests that the working alliance involving the young person seems to be more important than that of the therapist–parent alliance in predicting outcome in clinic-based therapy [18]. However, it remains to be seen whether results would be similar for a treatment that is fully online with minimal therapist assistance.

The present studies sought not only to determine whether a sound working alliance could be developed during a completely online CBT treatment for youth anxiety and whether it could predict treatment outcome, but also to examine the mechanism of action through which working alliance might influence outcome. As noted above, the mechanism through which working alliance is suggested to affect treatment outcome is through greater treatment compliance [19,20]. Studies examining the association between therapy compliance and treatment outcome have produced mixed findings, possibly reflecting the varying definitions and methods of measurement that have been employed [21]. Karver and colleagues [21], however, found that both alliance and compliance predicted outcome for CBT among depressed adolescents, although the effects were direct, rather than mediated. Chu and Kendall [19], in a study involving clinic-based CBT for anxious youth, found that high compliance was associated with significantly better therapy outcomes at posttreatment. Whether compliance mediated the relationship between alliance and compliance, however, was not investigated in that study.

The relationship between alliance, compliance, and outcome does not appear to have been examined for anxious youth when CBT is fully delivered online. We proposed that alliance, through its effect on compliance, would be particularly important in predicting outcome in online therapy for anxious young people. With online therapy, young people are required to work more independently, and therefore a strong working alliance is hypothesized to be particularly important for maintaining motivation, enthusiasm, and effort to complete therapy tasks. Furthermore, we proposed that a similar set of predictive factors would exist for parents who are also required to participate in online treatment and contribute significantly to the therapy process.

The present research therefore had two main aims. First, we compared the relative strength of working alliance in online versus clinic delivery of CBT for anxiety in youth to determine whether it would be feasible to establish a strong working alliance between young people (and their parents) and their online therapist, despite minimal contact and the absence of face-to-face interaction. Although we hypothesized that the attempts to promote a strong working alliance from the online therapy would lead to a strong working alliance among young people and their parents, we did not expect that the working alliance would be as strong for online as for clinic-based therapy, given the absence of face-to-face interaction and its associated advantages.

Second, we examined whether the strength of the working alliance between the young people and their online therapist predicted therapy outcome for anxious youth following online delivery of CBT. Furthermore, we examined whether the proposed association between working alliance and outcome would be mediated by compliance with therapy tasks. Given the important role of parents in CBT for child anxiety [8], we hypothesized an equivalent set of predictive relationships between working alliance, compliance, and child therapy outcome for parents who participated in the online therapy.

Given the recent report by McLeod [14] that the alliance–therapy outcome relationship was stronger among younger children than among adolescents for face-to-face therapy, we also conducted subsidiary analyses by age given that the sample ranged in age from 7 to 18 years. As the nature and importance of working alliance may differ between online and face-to-face therapy, we did not formulate any firm directional hypotheses with respect to age. We also conducted subsidiary analyses to determine whether the effects differed by gender. Although McLeod did not find gender differences in the relationship between working alliance and outcome for face-to-face therapy, whether such differences exist in online treatment remains to be determined.

The research involved data from two overlapping samples and as such they are reported as separate studies.

Study 1: Comparison of Working Alliance for Clinic and Internet Delivery

Methods

Participants

Study 1 participants were 73 youths (45 female, 28 male) aged 12 to 18 years (mean 13.91, SD 1.56), plus one or more of their parents. We recruited families from metropolitan areas of Brisbane and Sydney, Australia, through referrals from guidance officers, mental health professionals, and self-referrals following media publicity. The majority of youths were born in Australia ($n = 64$, 88%) and lived with both biological parents ($n = 58$, 79%). Comparison of family incomes with the Australian Bureau of Statistics [22] census information indicated that, on average, participants came from middle- to high-income Australian families and were relatively well educated.

To participate, youths were required to have access to a computer and the Internet at home, to read at a minimum age level of 10 years, and to meet criteria for a principal diagnosis of separation anxiety disorder, generalized anxiety disorder, social phobia, or specific phobia. Referrals were excluded if the young person met criteria for a principal diagnosis of obsessive-compulsive disorder, oppositional defiant disorder, posttraumatic stress disorder, conduct disorder, dysthymia, or major depressive disorder with a clinician severity rating greater than 4 on the Anxiety Disorders Interview Schedule for Children and Parents (ADIS-C/P) [23], or had an intellectual handicap, learning disability, or pervasive developmental disorder. Referrals were also excluded if the young person was engaging in current self-harming behavior or was receiving treatment for anxiety elsewhere.

Youths and parents were administered the ADIS-C/P via telephone by a trained psychologist. All youths met the criteria outlined in the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, text revision (DSM-IV-TR) for a primary anxiety disorder as determined by the ADIS-C/P (separation anxiety disorder: $n = 12$, 16%; generalized anxiety disorder: $n = 34$, 47%; social phobia: $n = 22$, 30%; specific phobia: $n = 4$, 6%; panic and agoraphobia: $n = 1$, 1%). Most ($n = 68$, 93%) youths had a secondary anxiety diagnosis. Following assessment, participants were randomly assigned to either the face-to-face therapy condition ($n = 38$) or the Internet condition ($n = 35$). Participants included those from the online and clinic conditions of a randomized controlled trial reported by Spence et al [5]. That study had also included a waitlist control condition and did not include the data relating to working alliance that are reported here.

Treatment Conditions

Clinic-Based Treatment

All clinic condition treatment sessions were conducted face-to-face with a therapist and followed the BRAVE program, which targets social phobia, generalized anxiety disorder, separation anxiety disorder, and specific phobia [24]. The program comprised 10 youth sessions and five parent sessions, each 1 hour long, as well as booster sessions at 1 and 3 months following treatment. Standard CBT anxiety management strategies were used, including psychoeducation, relaxation training, recognition of the physiological symptoms of anxiety, cognitive strategies of coping self-talk and cognitive restructuring, graded exposure, problem solving, and self-reinforcement. Parent sessions also taught anxiety management skills, in addition to parenting strategies to empower parents to help their child implement anxiety management skills.

Clinic sessions were conducted by registered psychologists at The University of Queensland, Griffith University, and Macquarie University psychology clinics. All clinicians received 2 days of training in use of the manualized clinic program, followed by weekly supervision.

Internet-Based Treatment

Online treatment participants completed BRAVE for Teenagers–ONLINE [7]. The BRAVE interventions have been

described in detail elsewhere [4,25], and therefore the following contains only a brief overview. The content, length, and number of session activities in the Internet program replicate those of the clinic-based version. Sessions are designed to be engaging, interactive, and age appropriate. Eye-catching graphics, sounds, games, and quizzes are used to maintain the youths' level of interest. Information is presented through interactive exercises and is followed by quizzes that check for correct understanding and provide personalized corrective or positive feedback through pop-up messages. The content of the intervention is designed to meet the developmental and cognitive level of youths, with age-appropriate scenarios, examples, and activities (example situations include school exams, job interviews, dating, and oral presentations).

The program is (minimally) therapist assisted, rather than self-help. Each family is assigned an online therapist (BRAVE trainer) who monitors their progress through the program and provides brief email feedback following each session. At no stage did any participants have face-to-face contact with their therapist and all other contact (ie, email or phone) was minimal. Clinician contact was restricted to brief, weekly emails and a short, 15-minute midprogram telephone call to assist in exposure hierarchy development. Most other contact with the online therapist was computer generated. Client responses to all session and homework activities were stored in an administrator section of the program and could be viewed by the therapist to guide the content of the weekly email. In addition, automated computer-generated emails were sent on behalf of the online therapist to congratulate participants for completing sessions, and personalized emails were sent to provide feedback about responses to quiz tasks. Personalized automated reminder emails were sent to advise when the next session would be available for completion or to provide prompts if the session was not completed by the due date. The first session also included a picture of the therapist and some brief biographical information about him or her, to which the client responded by providing information about him- or herself.

Youth and parent sessions were designed to be completed independently; however, there was no stipulation that parents could not help their child complete online sessions.

Measures

Anxiety Disorders Interview Schedule for Children–Parent and Child Version

The parent (ADIS-P) and a child (ADIS-C) interview schedules [23] were administered separately by trained clinicians. The ADIS determines the presence of anxiety disorders and other common disorders in children and adolescents. For each diagnosis obtained, the interviewer assigns a clinician severity rating ranging from 0 (no interference) to 8 (extreme interference). A rating of 4 (moderate interference) or more is considered to indicate a clinically significant diagnosis, with higher clinician severity rating scores representing more severe presentations of a disorder. The primary diagnosis was the diagnosis that was deemed to cause the most interference for the participant.

Interassessor reliability for the current study was determined from a random sample of interviews taken from 15% ($n = 11$) of families. This was conducted by two trained interviewers who were blind to the original diagnoses. Interassessor reliability was high for ADIS diagnoses, with a kappa value of 1 for the primary diagnosis and a correlation of .96 for the ADIS combined severity ratings.

Working Alliance Inventory–Short Form

The Working Alliance Inventory–Short Form (WAI-S) [26] is a 12-item scale, which includes three subscales: goal (agreement on the goals of therapy); task (agreement on how these goals will be achieved), and bond (the bond between participant and therapist). Participants were required to rate on a 7-point Likert scale from 1 (not at all true) to 7 (very true) the extent to which they believed each item was true of their relationship with their BRAVE trainer. Responses to all items were averaged to provide an overall working alliance score ranging from 1 to 7. Higher scores indicate a stronger working alliance (scores from 1 to 3 indicate a poor or negative relationship, 4 indicates a neutral position, and 5 to 7 indicate a good or positive therapeutic relationship). We obtained working alliance measures in both treatment conditions individually from youth participants and parents after they completed session 3. Working alliance was sampled early in treatment to minimize the potential confound between alliance scores and symptom improvement over the course of therapy [27].

Study 1 Procedure

For the purpose of the study, we created separate versions of the WAI-S for youths and parents. We adjusted the wording of the items to account for both the clinic and online delivery of treatment, the reading level of younger clients (in the youth version), and the focus on the youth's presenting problem in the parent version. The WAI-S has shown strong internal reliability [26,28,29] and factor structure in line with its theoretical model [30]. The WAI-S has also demonstrated adequate reliability and validity with youth populations [12,18].

The WAI-S was completed online by youths and parents, individually. Before using the measure in the present study, we examined whether the factor structure held up when administered online and in relation to therapist-assisted, online treatment. Consistent with prior research with adults in clinic-based therapy, it was predicted that confirmatory factor analysis would support a 3-correlated-factor model in which WAI-S items would load on intercorrelated factors relating to task agreement, goal agreement, and bond, with the covariation being explained to a large degree by a higher-order factor of working alliance in general [31]. The modified WAI-S was pilot tested with 137 clinically anxious children and adolescents (74 female, 63 male) aged between 7 and 18 years, plus one or more of their parents, participating in an online CBT intervention for anxiety. The results supported a single-factor model of working alliance for the young clients, with the comparative fit index = .96 and Tucker-Lewis Index = .97, demonstrating a good fit of the model (scores $\geq .95$ indicate strong fit), and root mean square error approximation = 0.075 (90% confidence interval 0.05–0.10) and standardized root mean square residual = .03 (scores equal to or greater than .08 and .05, respectively,

represent good data fit.). There were significantly correlated error terms between some items. A 3-correlated-factor model or higher-order model did not add significantly to the fit of the 1-factor model. Cronbach alpha for the youth WAI-S was .96. For parents, the data were well explained by a 2-factor model (bond and task/goal combined), loading onto a higher-order factor of working alliance, with comparative fit index = .95, Tucker-Lewis Index = .94, root mean square error approximation = 0.087 (90% confidence interval 0.06–0.11), and standardized root mean square residual = .07. Internal consistency was .94 for the parent WAI-S. Taken together, the results supported the use of the WAI-S total scores for both parents and youths in the online delivery of CBT for anxiety disorders and justified its use in the present study.

Results

Pretreatment Comparison

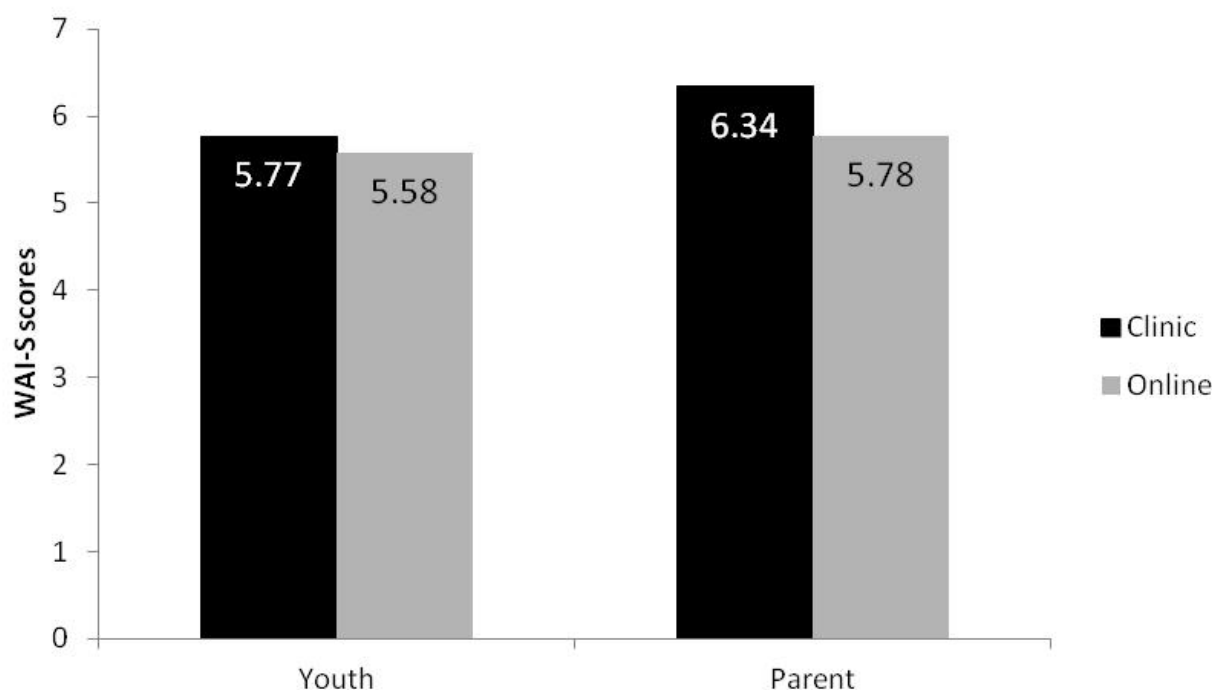
Chi-square analyses showed no significant differences between the online and clinic conditions on gender ($n = 73$, $\chi^2_1 = 1.6$, $P = .24$), parental marital status ($n = 73$, $\chi^2_5 = 6.1$, $P = .30$), or child living situation ($n = 73$, $\chi^2_3 = 0.2$, $P = .98$). Multivariate analysis of variance also revealed there were no significant differences between the clinic and online conditions for youth's age, mother's age, father's age, education level of mother and father, and combined family income (Pillai $F_{6,66} = 2.04$, $P = .07$, $\eta_p^2 = .16$). Chi-square analyses also showed no significant

differences between the clinic and online conditions for the type of primary anxiety diagnoses ($n = 73$, $\chi^2_5 = 2.5$, $P = .78$) or the number of comorbid nonanxiety disorders ($n = 73$, $\chi^2_1 = 0.3$, $P = .58$). A multivariate analysis of variance revealed no significant differences in the severity of the primary anxiety diagnosis or the total number of anxiety diagnoses received (Pillai $F_{2,71} = 0.24$, $P = .87$, $\eta_p^2 = .01$).

WAI-S Scores for Youths and Parents by Therapy Condition

Youth WAI-S scores ranged from 2.42 to 7.00 (mean 5.77, SD 1.20) for the clinic condition, and from 2.08 to 7.00 (mean 5.58, SD 1.34) for the online condition. There were no significant differences between the groups ($F_{1,73} = 0.44$, $P = .51$, $\eta_p^2 = .006$, Cohen $d = 0.15$), and youths in both groups reported strong working alliance (see Figure 1). Clinic condition parent WAI-S scores ranged from 4.00 to 7.00 (mean 6.34, SD 0.67), and online condition parent WAI-S scores ranged from 3.17 to 7.00 (mean 5.78, SD 1.06). Although both online and clinic condition parents rated the working alliance with their therapist as positive and strong, those in the clinic condition rated alliance with their face-to-face therapist significantly higher than parents in the online condition ($F_{1,70} = 6.76$, $P = .01$, $\eta_p^2 = .09$, Cohen $d = 0.64$). No significant correlations were found between youth and parent WAI-S scores in either the clinic ($r = .09$, $P = .60$) or online ($r = .20$, $P = .22$) treatment conditions.

Figure 1. Mean Working Alliance Inventory-Short Form (WAI-S) scores for youths and parents in the face-to-face condition (clinic) and Internet condition (online).



Study 2: Prediction of Online Therapy Outcome From Working Alliance and Compliance

Methods

Participants

Participants in study 2 were 132 children and adolescents who met criteria for an anxiety disorder plus at least one parent, who had completed a minimum of three sessions of online-delivered CBT (the point at which the working alliance measure was administered) and for whom outcome data were obtained at 6-month follow-up. Participants in study 2 included the online group from study 1, but we supplemented the sample to include additional adolescents and expanded the age range to include youth aged from 7 to 12 years. This extended sample increased the statistical power to enable examination of predictive factors using regression analyses and the testing of moderating variables such as age and gender. Given the extended age range and addition of further participants to those reported in study 1, the participant characteristics for study 2 will be reported in full. There were 70 female and 62 male participants, aged 7 to 18 years (mean 12.12, SD 2.50), plus one or more of their parents. The majority were born in Australia ($n = 119$, 90.1%) and were living with both biological parents (107, 81.1%). On average, the participants came from middle- to high-income Australian families and were relatively well educated. The inclusion and exclusion criteria were identical to those in study 1. All children and adolescents met DSM-IV-TR criteria for a primary anxiety disorder, including separation anxiety disorder ($n = 28$, 21%), generalized anxiety disorder ($n = 53$, 40%), social phobia ($n = 41$, 31%), or specific phobia ($n = 10$, 8%), according to the ADIS-C/P. The mean clinician severity rating for the principal anxiety diagnosis was 5.95 (SD 0.76), and the mean number of anxiety disorders was 2.87 (SD 1.44). In total, 117 participants (88.6%) had at least two diagnoses.

Study 2 Procedure

Procedures were identical to those outlined for study 1, with the exception that participants were restricted to those participating in online therapy (with no clinic comparison group). Outcome was evaluated 6 months after the 12-week treatment period. The follow-up assessment included an ADIS interview completed with each child or adolescent and one parent, and the clinician-rated Children's Global Assessment Scale (CGAS) [32]. The follow-up interviews were completed by trained clinicians who were blind to participants' treatment condition. Treatment effectiveness was measured at 6-month follow-up to provide a stable and longer-term indicator of outcome, with the key indicator for assessing the prediction of outcome being the CGAS rating. We selected the CGAS rather than clinician ratings of primary anxiety disorder severity (clinician severity rating) or parent/youth questionnaire measures, as it provided a more comprehensive indication of overall improvement made by the child or adolescent, rather than improvement made for only the principal anxiety disorder or a subjective response to symptom checklists.

Measures

The WAI-S and the ADIS-C/P described for study 1 were also used in study 2. We administered additional youth and parent self-report measures of anxiety in study 2 to determine whether youths improved from pretreatment to 6-month follow-up after completing the online CBT intervention, and whether there was variability in outcome in order to justify the examination of predictors.

Primary Outcome Measure

Children's Global Assessment Scale

The CGAS [32] provides a single global rating of functioning, assigned to the youth by the independent interviewing clinician, on a scale of 0 to 100, where lower scores indicate poorer functioning. A rating is given based on the child's or adolescent's most impaired level of general functioning for the specified time period by selecting the lowest level that describes his or her functioning on a hypothetical continuum of health-illness, benchmarked against anchor points in a descriptive glossary. The CGAS has demonstrated sufficient sensitivity in detecting treatment outcomes for child anxiety (eg, March et al [4]). It has shown high interrater reliability with a correlation ranging from .59 to .90 [32-35].

Interassessor reliability for the current study was determined from a random sample of audiotaped interviews taken from 15% ($n = 20$) of families, conducted by two trained, independent interviewers who were blind to the original diagnoses. Interassessor reliability was high for the CGAS ratings ($r = .94$).

Secondary Outcome Measures

Child Behavior Checklist-Revised

Parents completed the Child Behavior Checklist (CBCL) [36], a psychometrically sound measure of behavioral problems for young people aged between 4 and 18 years. Parents were asked to rate whether an item applied to their child, using a 3-point scale (0 = not true; 1 = somewhat or sometimes true; 2 = very true or often true). We used only the internalizing scale in the present study.

Spence Children's Anxiety Scale-Parent Version and Child Version

We used the Spence Children's Anxiety Scale-Parent Version (SCAS-P) and Child Version (SCAS-C) [37,38] to measure child and adolescent anxiety symptoms. The SCAS-P and SCAS-C assess symptoms relating to separation anxiety, social phobia, obsessive-compulsive disorder, panic-agoraphobia, generalized anxiety, and fears of physical injury according to symptom clusters represented by the DSM-IV-TR. Children and parents rate on a 4-point scale how frequently the child experiences a particular event, with higher scores representing higher levels of anxiety. We used only the total scores in the study's analyses. The SCAS-C has demonstrated sound psychometric properties, with an internal reliability coefficient of .92 for the total score, and its factor structure has been confirmed in several studies [37-39]. The parent version of the SCAS also demonstrates sound psychometric properties, with high internal consistency for total anxiety scores ($\alpha = .89$) [40].

Compliance With Internet Sessions

Program compliance was operationalized as the percentage of therapy sessions and homework tasks completed by the participant by the 6-month follow-up. As described above, online sessions comprised multiple activities. During each session, participants were required to respond to exercises and quizzes that involved providing responses or typing answers into the program. These responses were stored electronically in individual file logs, accessed through the administrative section of the online program. The total number of tasks to be completed included all session activities where the participant was able to enter a response into the program, from sessions 1 through 10 for youths and 1 through 6 for parents (or five sessions for parents of adolescents, as material from the six sessions was condensed into five sessions for parents of this age group).

Responses to tasks were scored as either 0 = not completed, or 1 = completed. For an item to be considered completed and to receive a score of 1, the participant must have attempted the activity with a meaningful response. A trained researcher reviewed each participant's log file to determine whether the participant had attempted to complete each task in a meaningful way and assigned a score of 0 or 1 accordingly. The total number of tasks completed was divided by the total number of tasks for the entire intervention to obtain a percentage compliance score.

A second trained researcher selected a random sample of 15% ($n = 20$) of log files and coded them using the same guidelines for scoring. Independent interrater agreement on total compliance scores was high ($\kappa = 0.975$).

Internet Therapy Program

The online intervention in the current study was outlined above for study 1. However, given the age range of the young people involved, we created two versions of the site so that the reading age required, illustrations, examples, and graphics were developmentally appropriate. For example, the site for the younger children used a cartoon character (Brave Buddy) who represented the coping model, whereas the adolescent version used cartoon teenaged coping models to illustrate the concepts.

Analyses

We conducted a series of standard and hierarchical multiple regression analyses to test whether the variables of interest were associated with treatment outcome or treatment compliance. Analyses to test mediation were also used where appropriate as outlined by Baron and Kenny [41]. Baseline CGAS scores were entered into each hierarchical equation at step 1 of all analyses in the prediction of CGAS at 6-month follow-up, to control for pretreatment severity.

Results

Treatment Outcome

The first step was to determine whether changes in anxiety over time were significant, with sufficient variation in outcome to justify examining prediction of response to treatment. As these analyses were not central to the purpose of the study, we discuss them only briefly.

Changes in Anxiety Following Treatment

Table 1 and Table 2 present data regarding treatment outcome and analyses. At 6-month follow-up, significant improvements were evident for all outcome measures, including diagnostic criteria, overall functioning on the CGAS, and reductions in internalizing and anxiety symptoms over time for CBCL, SCAS-C, and SCAS-P. Taken together, the results indicated significant improvements in emotional wellbeing over time, but with variation in response, justifying the examination of predictors of treatment outcome.

Compliance With Internet-Based Treatment Tasks

The mean number of sessions recorded as completed by children and adolescents was 8.86 (SD 1.90) out of 10 sessions. Parents completed an average of 5.74 sessions (SD 0.66) out of 6 in the child program and 4.76 sessions (SD 0.56) out of 5 in the adolescent program. In terms of session tasks, youths completed an average of 85% (SD 20.22, range 26% to 100%, 25 to 95 tasks) of available tasks and parents completed an average of 89% (SD 12.32, range 27% to 100%, 15 to 55 tasks) of tasks available to them.

Working Alliance

The mean child and adolescent WAI-S at the end of session 3 was 5.85 (SD 1.09) with a range from 2.08 to 7.00, and the mean parent WAI-S was 6.07 (SD 0.76) with a range from 2.92 to 7.00. On average, both parents and youths formed a strong working alliance with their online therapist.

Correlations Between Alliance, Compliance, and Outcome Measures

There was a significant and positive correlation between youth alliance and parent alliance ($r = .31$, $P < .001$), such that higher levels of parent alliance were associated with higher levels of youth alliance. This finding differed from that of study 1, in which youth alliance and parent alliance were not significantly correlated. In study 2, this can be explained by a much higher correlation between parent and child WAI-S scores among 7- to 11-year-old children ($r = .51$, $P < .001$) than among the adolescents in the sample ($r = .26$, $P < .001$).

Table 1. Treatment outcomes for children and adolescents with at least one anxiety disorder assessed in study 2.

Measure	Pretreatment		6-month follow-up		<i>F</i> value	df (hypothesis, error)	η_p^2
	Mean	SD	Mean	SD			
CSR ^a	5.96	0.76	2.44	1.96	413.17***	1,130	.76
CGAS ^b	49.82	4.95	70.80	11.09	554.93***	1,130	.81
Number of diagnoses	2.87	1.44	0.82	1.16	265.34***	1,130	.67
Youth SCAS-C ^c	39.59	17.14	22.04	13.63	177.34***	1,112	.61
Parent SCAS-P ^d	32.33	14.04	18.01	8.49	139.85***	1,116	.55
CBCL ^e	21.19	7.96	10.87	7.48	146.99***	1,117	.60

^a Clinician severity rating.^b Children's Global Assessment Scale.^c Spence Children's Anxiety Scale-Child Version.^d Spence Children's Anxiety Scale-Parent Version.^e Child Behavior Checklist.**Table 2.** Children and adolescents free of primary diagnosis and any diagnosis at pretreatment and 6-month follow-up.

Type of diagnosis	Pretreatment		6-month follow-up		χ^2	df
	n	%	n	%		
Free of primary diagnosis	132	0	87	66.4	85.01***	1,131
Free of any diagnosis	132	0	70	53.4	68.01***	1,131

*** $P < .001$.

A positive relationship was also found between parent compliance and youth compliance ($r = .53$, $P < .001$). Youth alliance measured after session 3 was significantly and positively associated with youth compliance by 6-month follow-up ($r = .30$, $P < .001$). Parent alliance was not significantly associated with parent compliance or youth compliance. Youth alliance and parent alliance, and youth compliance and parent compliance were not significantly correlated with the outcome (CGAS at 6-month follow-up). As expected, baseline CGAS was significantly correlated with CGAS at follow-up ($r = .40$, $P < .001$).

Demographic variables (age of youth, household income, and mother's and father's education) were not significantly correlated with any of the independent variables (parent alliance or youth alliance, and parent compliance or youth compliance), the baseline measure of severity (CGAS), or the dependent variable (CGAS at 6-month follow-up), except for mother's education, which was positively and significantly correlated with youth alliance ($r = .21$, $P < .05$).

Alliance Predicting Outcome

To examine whether alliance predicted outcome, we performed a hierarchical multiple regression. Pretreatment CGAS scores were entered on the first step to control for the effects of baseline severity. Youth alliance was entered on the second step. The overall model was significant at step 2 ($F_{2,128} = 12.22$, $P < .001$). Baseline CGAS was a significant predictor of CGAS at 6-month follow-up (beta = .40, $t_{129} = 4.93$, $P < .001$). When we examined the unique effects of youth alliance at step 2, it was not a

significant predictor of CGAS at 6-month follow-up (beta = .08, $t_{129} = 0.95$, $P = .34$). Similarly, when we examined the unique effects of parent alliance at step 2, it was not a significant predictor of CGAS at 6-month follow-up (beta = .12, $t_{129} = 1.49$, $P = .14$).

Alliance Predicting Compliance

The analysis revealed that youth alliance significantly and positively predicted youth compliance (beta = .30, $F_{1,127} = 12.32$, $P = .001$) and accounted for approximately 9% of the variance in compliance scores. However, parent alliance did not significantly predict parent compliance ($F_{1,126} = 0.36$, $P = .55$).

Compliance Predicting Outcome

Pretreatment CGAS scores were entered on the first step to control for the effects of baseline severity and youth compliance entered on the second step. The overall model was significant ($F_{2,128} = 13.76$, $P < .001$). When we examined the unique effects of youth compliance at step 2, it was not a significant predictor of CGAS at 6-month follow-up (beta = .08, $t_{129} = 0.95$, $P = .34$). Similarly, when we examined the unique effects of parent compliance at step 2, it did not significantly predict CGAS at 6-month follow-up (beta = .14, $t_{129} = 1.74$, $P = .08$). Given the absence of direct effects between alliance and compliance with outcome, it was not therefore appropriate to examine a mediation model.

Gender as a Moderator

We examined gender as a possible moderating variable in the relationship between youth alliance and youth compliance with outcome at 6-month follow-up. A moderated hierarchical regression analysis was performed with baseline CGAS entered first to control for pretreatment severity. Youth alliance and gender were entered into the regression at step 2. The interaction term was entered on step 3. However, the interaction between youth alliance and gender did not significantly predict outcome ($\beta = .01$, $t_{129} = .16$, $P = .87$). We then performed a moderated hierarchical regression analysis for the interaction between youth compliance (centered) and gender in predicting outcome, but this was also not significant ($\beta = .09$, $t_{129} = 1.11$, $P = .27$).

Age as a Moderator

Youth age was then examined as a moderating variable in the prediction of CGAS outcome, based on age as a continuous variable. The centered variables, youth alliance and age, were entered into the regression analysis at step 2 after controlling for baseline CGAS. We found a main effect for age ($\beta = .17$, $t_{129} = 2.01$, $P = .05$), such that older age was associated with significantly better outcome on CGAS at 6 months. At step 3, the interaction term for the two variables was added. The overall model was significant ($F_{4,123} = 8.54$, $P < .001$). There was no main effect for youth alliance ($\beta = .08$, $t_{129} = .97$, $P = .33$), and the main effect for age remained significant. The interaction between youth alliance and age significantly predicted CGAS outcome ($\beta = .17$, $t_{129} = 2.08$, $P = .04$), which accounted for 3% of the variance in outcome.

We conducted simple slopes analyses for younger (1 SD below the mean) and older (1 SD above the mean) youths, and low (1 SD below the mean) and high (1 SD above the mean) alliance, to aid interpretation. Simple slopes analyses (see Figure 2) indicated that for adolescents there was a significant and positive relationship between alliance and change in CGAS scores from pretreatment to 6-month follow-up, such that a stronger alliance was associated with a greater positive change in CGAS and thus a better treatment outcome ($B = 2.73$, $SE = 1.15$, $t_{128} = 2.37$, $P = .02$). For younger youths, the relationship between alliance and change in CGAS scores from pretreatment to 6-month follow-up was not significant ($B = -1.11$, $SE = 1.33$, $t_{128} = -0.83$, $P = .41$).

Change in CGAS scores aided in the interpretation of the simple slopes analyses and figures.

We then tested age as a moderating variable in the relationship between parent alliance and outcome at 6-month follow-up. The centered variables of parent alliance and age of youth were entered into the regression analysis at step 2, after controlling for baseline CGAS at step 1, followed by the interaction term for the two centered variables at step 3. The overall model was

significant ($F_{4,121} = 8.96$, $P < .001$). There was no significant main effect for parent alliance ($\beta = .11$, $t_{125} = 1.40$, $P = .17$) or age ($\beta = .16$, $t_{125} = 1.93$, $P = .06$). However, there was a significant interaction between parent alliance and age of youth ($\beta = .19$, $t_{125} = 2.36$, $P = .02$) that accounted for 4% of the variance.

Simple slopes analyses (see Figure 3) indicated that for older youths there was a significant and positive relationship between parent alliance and change in CGAS scores from pretreatment to 6-month follow-up, such that a stronger parent alliance was associated with a greater positive change in CGAS and thus a better treatment outcome ($B = 4.11$, $SE = 1.55$, $t_{125} = 2.65$, $P < .01$). For younger youths, the relationship between parent alliance and change in CGAS scores from pretreatment to 6-month follow-up was not significant ($B = -0.86$, $SE = 1.59$, $t_{125} = -0.54$, $P = .59$).

We then tested age as a moderating variable in the relationship between youth compliance and outcome at 6-month follow-up. The centered variables of youth compliance and youth age were entered into the regression analysis at step 2, after controlling for baseline CGAS at step 1, followed by the interaction term for the two centered variables at step 3. The overall model was significant ($F_{4,126} = 10.03$, $P < .001$). There was no significant main effect for youth compliance ($\beta = .06$, $t_{128} = 0.75$, $P = .46$) or age ($\beta = .08$, $t_{128} = 0.92$, $P = .36$). However, there was a significant interaction between youth compliance and age ($\beta = .25$, $t_{128} = 3.02$, $P = .003$) that accounted for 6% of the variance.

Simple slopes analyses (see Figure 4) indicated that for older youths there was a significant and positive relationship between youth compliance and change in CGAS scores from pretreatment to 6-month follow-up, such that a higher youth compliance was associated with a better treatment outcome ($B = 1.79$, $SE = 0.50$, $t_{125} = 3.55$, $P < .001$). For younger youths, the relationship between compliance and change in CGAS scores from pretreatment to 6-month follow-up was not significant ($B = -0.48$, $SE = 0.52$, $t_{125} = -0.93$, $P = .35$).

Lastly, we tested age as a moderating variable in the relationship between parent compliance and outcome on the CGAS at 6-month follow-up. The centered variables of parent compliance and age of youth were entered into the regression analysis at step 2, after controlling for baseline CGAS at step 1, followed by the interaction term for the two centered variables at step 3. The overall model was significant ($F_{4,126} = 8.46$, $P < .001$). There was no significant main effect for parent compliance ($\beta = .14$, $t_{128} = 1.70$, $P = .09$) or age ($\beta = .08$, $t_{128} = 0.87$, $P = .39$). The interaction between parent compliance and age did not significantly predict outcome ($\beta = .14$, $t_{128} = 1.63$, $P = .11$).

Figure 2. Conditional effects of youth age on the relationship between youth working alliance and change in child global functioning, as measured by the Children's Global Assessment Scale (CGAS), from pretreatment to 6-month follow-up. Younger/older age represents ± 1 SD (2.5 years) below/above the mean (12.12 years).

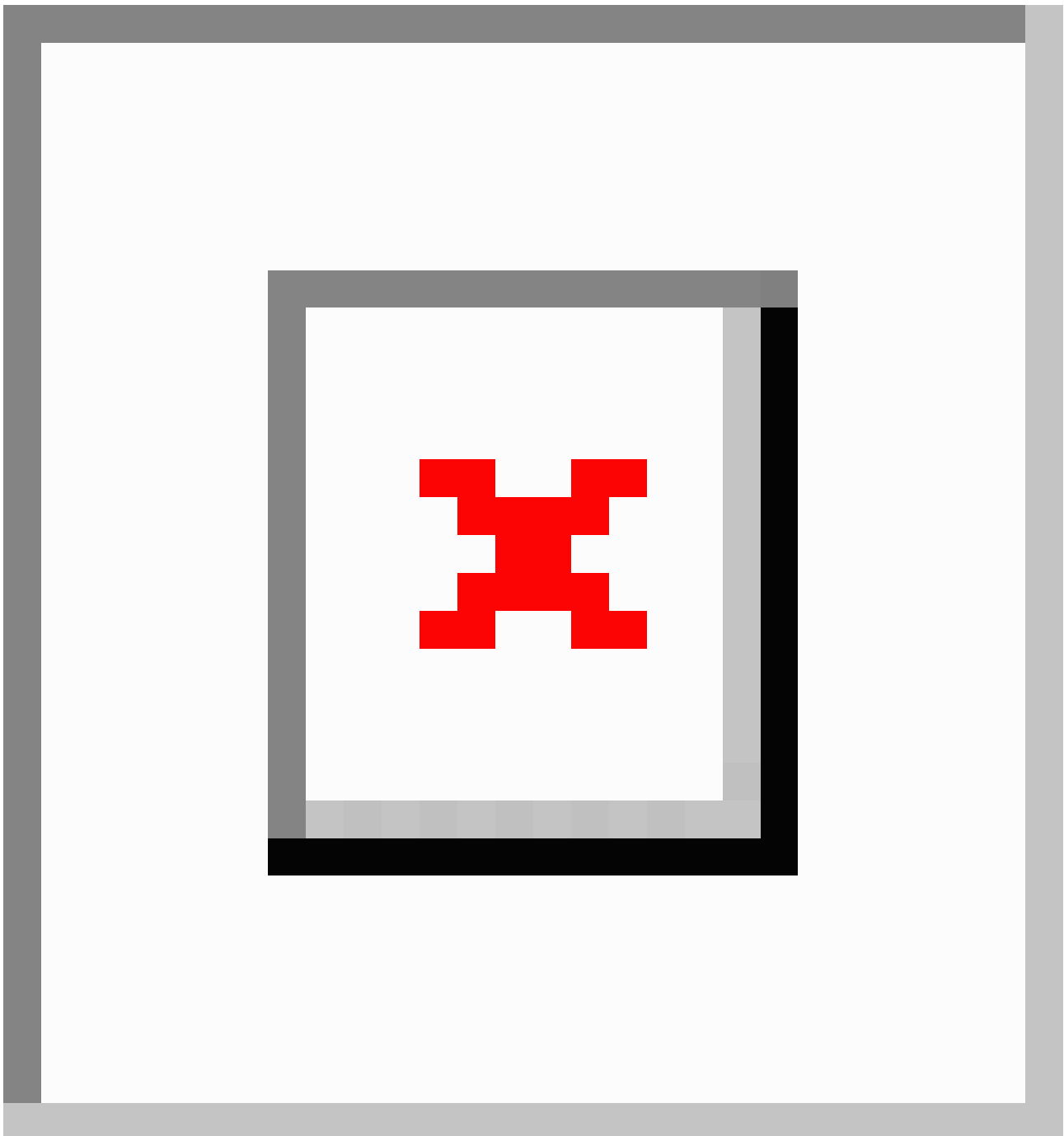


Figure 3. Conditional effects of youth age on the relationship between parent working alliance and change in child global functioning, as measured by the Children's Global Assessment Scale (CGAS), from pretreatment to 6-month follow-up. Younger/older age represents ± 1 SD (2.5 years) below/above the mean (12.12 years).

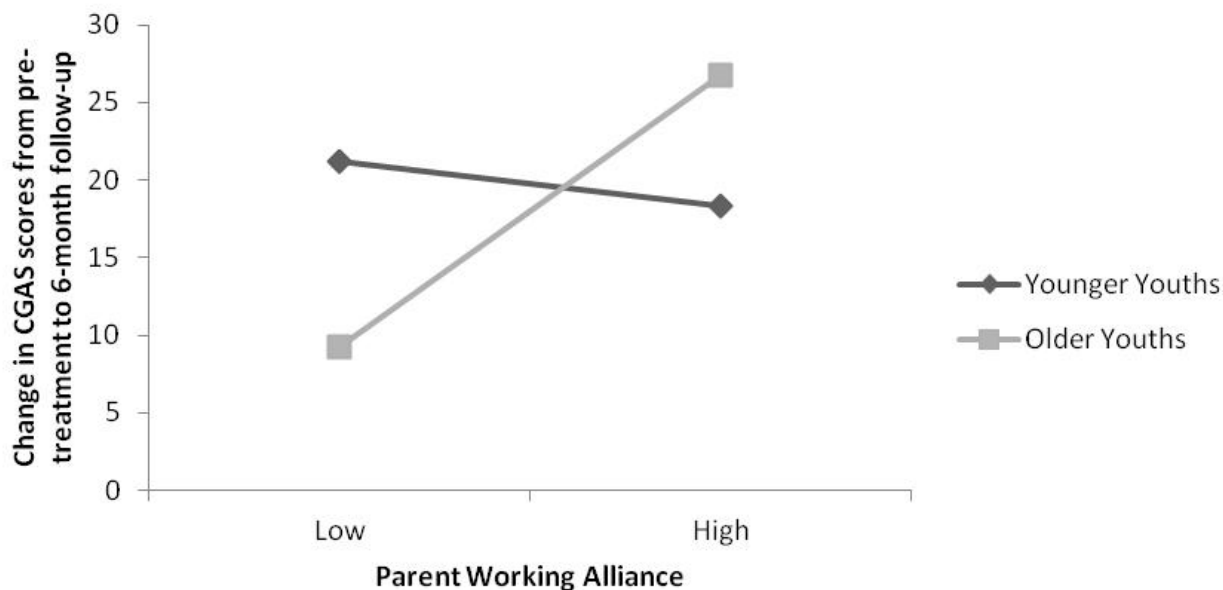
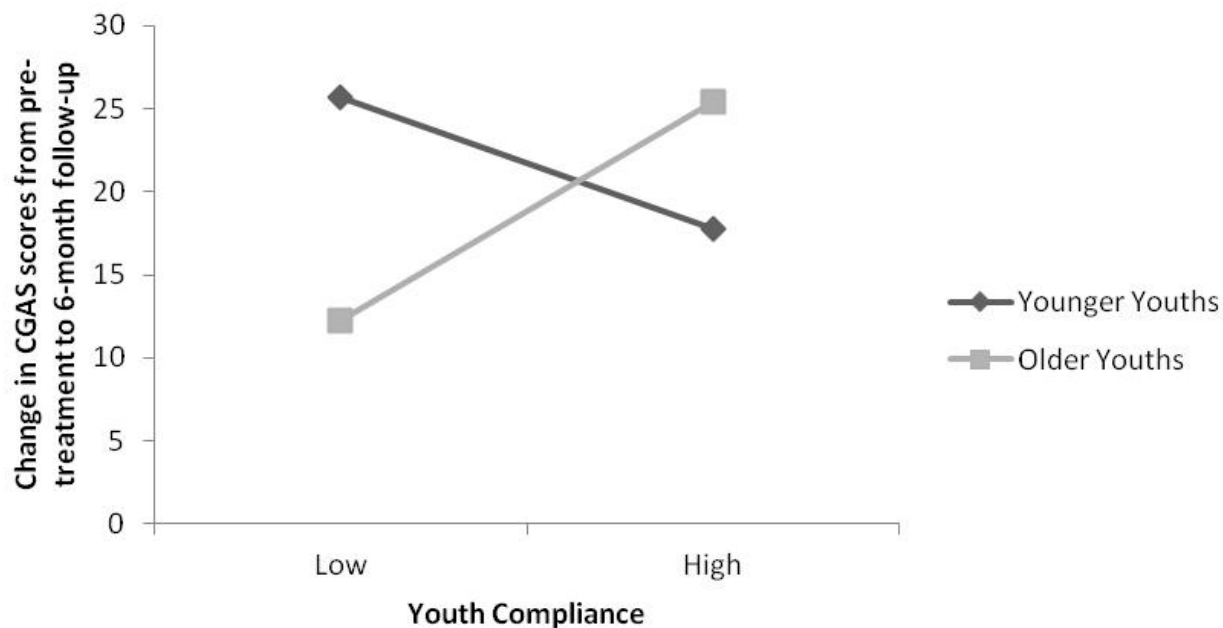


Figure 4. Conditional effects of youth age on the relationship between youth compliance and change in child global functioning, as measured by the Children's Global Assessment Scale (CGAS), from pretreatment to 6-month follow-up. Younger/older age represents ± 1 SD (2.5 years) below/above the mean (12.12 years).



Age Differences Between Variables

To inform possible explanations for the differential pattern of results by age, we divided the sample into younger children (aged 7 to 11 years) and adolescents (aged 12 to 18 years). Table 3 presents the means, standard deviations, and regression analyses for adolescents and children. Adolescents had significantly lower CGAS scores at baseline than the younger group ($F_{1,130} = 8.05$, $P < .01$, $\eta_p^2 = .06$) but no significant differences in CGAS at 6-month follow-up ($F_{1,129} = 0.02$, $P =$

$.90$, $\eta_p^2 = .00$). Working alliance at session 3 was slightly but significantly lower among adolescents than among children ($F_{1,128} = 4.01$, $P = .05$, $\eta_p^2 = .03$).

Parents' ratings of alliance measured after session 3 were not significantly different for the two age groups ($F_{1,126} = 0.22$, $P = .64$, $\eta_p^2 = .00$). However, parents of adolescents were significantly more compliant with the treatment program by

6-month follow-up than were parents of children ($F_{1,131} = 6.87$, $P = .01$, $\eta_p^2 = .05$).

Table 3. Means, standard deviations, and regression analyses for alliance, compliance, and CGAS scores compared between children and adolescents.

Measure	Children (7–11 years)		Adolescents (12–18 years)		<i>F</i>	df	<i>P</i> value	η_p^2
	Mean	SD	Mean	SD				
Youth alliance	6.10	.80	5.71	1.21	4.01	1,128	.05	.03
Parent alliance	6.06	.76	6.07	.77	0.22	1,126	.64	.00
Youth compliance	84.09	20.98	84.89	19.90	0.05	1,131	.83	.00
Parent compliance	85.00	16.11	91.68	12.72	6.87	1,131	.01	.05
Baseline CGAS ^a	51.47	4.77	48.98	4.86	8.05	1,130	.01	.06
Follow-up CGAS	70.64	10.79	70.89	11.32	0.02	1,129	.90	.00
Δ CGAS ^b	19.17	10.27	22.00	10.07	2.35	1,131	.13	.02

^a Children's Global Assessment Scale.

^b Change in CGAS from pretreatment to 6-month follow-up.

Examination of Mediation Model for Adolescents

Given the findings indicating that age significantly moderated the relationships between youth alliance and outcome and between youth compliance and outcome, and that compliance and outcome and alliance and outcome were significantly associated in the simple slopes analyses for older youths (aged 12 to 18 years), we considered it appropriate to examine the mediation model with the older age group separately. Table 4 summarizes the results.

First, youth alliance significantly predicted CGAS at 6-month follow-up ($\beta = .22$, $t_{79} = 2.21$, $P = .03$) and accounted for 5% of the variance. Second, compliance significantly predicted CGAS at 6-month follow-up ($\beta = .22$, $t_{84} = 2.22$, $P = .03$) and accounted for 5% of the variance. Third, adolescent alliance significantly and positively predicted adolescent compliance ($\beta = .38$, $F_{1,80} = 13.10$, $P = .01$), accounting for approximately 14% of the variance in compliance scores.

In the final step of testing, the mediation model examined whether the significant relationship between adolescent alliance and outcome decreased in significance or became nonsignificant when adolescent compliance was entered into the equation. Pretreatment CGAS scores were entered at the first step to control for the effects of baseline severity. Adolescent compliance and adolescent alliance were entered at the second step. The overall model was significant ($F_{3,77} = 9.84$, $P < .001$); however, when adolescent compliance was controlled for, adolescent alliance no longer significantly predicted CGAS scores at 6-month follow-up ($\beta = .17$, $t_{79} = 1.56$, $P = .12$). Because adolescent compliance was not significant at step 2, despite adolescent alliance becoming nonsignificant, a mediation model was not supported. Partial mediation effects were also tested. Bootstrapping results for the indirect effect [42] confirmed that no mediation or partial mediation existed, with the 95% bias-corrected bootstrap confidence interval based on 5000 bootstrap samples (−1.12 to 1.40, with a point estimate of 0.09).

Table 4. Summary of mediational analyses for the adolescent group.

Group	R^2	Adjusted R^2	$R^2 \Delta$	$F \Delta$	df	Beta	sr^2
Adolescent alliance predicting adolescent compliance at 6-month follow-up							
Step 1							
Adolescent alliance	.14	.13	.14	13.10	1,80	.38**	.14
Adolescent alliance predicting CGAS^a scores at 6-month follow-up							
Step 1							
Baseline CGAS scores	.21	.20	.21	20.89	1,79	.51***	.24
Step 2							
Baseline CGAS scores	.26	.24	.05	4.89	2,78	.50***	.24
Adolescent alliance						.22*	.05
Adolescent compliance predicting CGAS scores at 6-month follow-up							
Step 1							
Baseline CGAS scores	.21	.20	.21	21.68	1,82	.49***	.24
Step 2							
Baseline CGAS scores	.25	.24	.05	4.91	2,81	.48***	.24
Adolescent compliance						.22*	.05
Adolescent compliance as a mediator in the relationship between adolescent alliance and outcome on CGAS at 6-month follow-up							
Step 1							
Baseline CGAS scores	.25	.24	.25	13.30	1,79	.52***	.26
Step 2							
Baseline CGAS scores	.28	.25	.02	2.42	3,77	.51***	.26
Adolescent compliance						.16	.02
Adolescent alliance						.17	.02

^a Children's Global Assessment Scale.

* $P < .05$; ** $P < .01$; *** $P < .001$.

Thus, for the adolescents, although both alliance and compliance significantly predicted treatment outcome at follow-up when entered separately, and alliance significantly predicted compliance, compliance was not found to significantly mediate the relationship between adolescent alliance and CGAS scores at 6-month follow-up. Rather, both alliance and compliance had direct effects on treatment outcome.

Discussion

The present research confirms that a strong working alliance can be established for both parents and their children during online therapy, despite the absence of face-to-face contact and with only minimal therapist contact through email and a brief phone call. Indeed, from the perspective of adolescents, there was no significant difference in the strength of the working alliance between online and clinic delivery of CBT for youth anxiety. Although parents reported a significantly greater working alliance from clinic-based than from online therapy, the difference was extremely small and the working alliance in online therapy was still strong. Firm conclusions cannot be drawn about the determinants of the strong working alliance; however, it seems likely that the personalized pop-ups,

automated feedback, and emails were beneficial in achieving this goal. The research also demonstrated the construct validity of the WAI-S as an indicator of the quality of therapist–client relationship in minimal therapist-assisted online treatment for young people and their parents and the strong internal reliability of the measure. The strong psychometric properties justified its use in the subsequent analyses.

The positive treatment response demonstrated in study 2 following online therapy was consistent with previous research involving either partial or full Internet or computer delivery of CBT interventions with young people [3,4,7,25]. In study 2, 66% ($n = 87$) of treated children and adolescents were free from their primary anxiety diagnosis by 6-month follow-up, which is consistent with much of the clinic-based literature in this area [6]. However, also consistent with clinic-based treatment of youth anxiety disorders was the finding that a considerable proportion of youth retained at least one clinical diagnosis of anxiety after treatment ($n = 62$, 47%). It was important, therefore, to investigate factors that may predict treatment outcome and that could potentially be influenced in order to produce better outcomes.

Although working alliance and therapy compliance were not found to significantly predict treatment outcome for the total sample, a consistently different pattern of results emerged according to the age of the young person. For younger children, there were no significant relationships between working alliance, compliance, and outcome. In contrast, for adolescents a stronger working alliance for both the youth and the parent predicted greater positive change in CGAS scores by 6-month follow-up. Similarly, higher youth compliance predicted more positive changes in CGAS scores from pretreatment to 6-month follow-up for adolescents. Also, consistent with our hypotheses, greater working alliance among adolescents predicted greater therapy compliance. However, contrary to predictions, compliance was not found to mediate the alliance–outcome relationship, with the effects being direct rather than mediated. Interestingly, this finding is consistent with that of Karver and colleagues [21], who found that both alliance and compliance predicted outcome for CBT among depressed adolescents, and that the effects were direct, rather than mediated. Whether mediated or not, the results emphasize the importance of a strong working alliance and compliance with therapy tasks in leading to better outcomes in the adolescent group.

It is important to consider why positive associations between alliance, compliance, and outcome should be evident for adolescents but not for the younger age group. Both age groups, in general, formed a strong working alliance with their online therapist, and compliance was equivalent across age groups. Outcome was slightly stronger among the adolescent group, although this may reflect their slightly poorer level of adjustment at baseline. The parents of the adolescents also showed slightly, but significantly, higher levels of therapy compliance than the parents of the younger children. One can only speculate about some of the reasons why working alliance and compliance were more important in predicting outcome for adolescents. Perhaps with younger children, anxiety problems are more transient and the rate of improvement may be faster, leading to earlier cessation of therapy tasks because improvements are already evident. Future research needs to track symptom severity more frequently during therapy in order to examine rates of change, as it is likely that the relationship between compliance and outcome is a complex one and there may be many reasons why people do not comply with treatment tasks. Another point to consider is that the nature of the relationship between young people changes with age. Younger children may be more reliant on their parents for support during therapy, and the quality of their perceived relationship with their online therapist may be less important to determining their level of treatment compliance or outcome. Adolescents, on the other hand, tend to be more independent from their parents and may place greater value on

their online therapist for support and encouragement. This proposition is consistent with the finding of little correlation between parent and adolescent WAI-S scores with their therapist, whereas the association between younger children's and their parents' WAI-S scores was much stronger. We must also question whether younger children were as free to express their personal views of the WAI-S or were influenced in their judgments by parental guidance. Perhaps their understanding or view of alliance is different from that of adolescents. Clearly, the findings emphasize how important it is to consider age differences in this type of research. The sample size in study 2 was sufficiently large to enable examination of age effects, something that has not been possible in studies with smaller sample sizes and insufficient power.

Several limitations must be considered for the current studies. First, it is possible that the inclusion of both qualitative *and* quantitative measures of compliance may have provided a more comprehensive assessment of compliance than the quantitative-only analyses in the current studies. It is noteworthy, however, that for adolescents, compliance significantly predicted CGAS at 6-month follow-up, suggesting that more sensitive qualitative measurements of compliance may be most relevant in younger cohorts and their parents. Second was the inability to examine in detail how working alliance, compliance, and anxiety changed over time on a session-by-session basis. This means that it is not possible to separate cause–effect relationships. Finally, it is not clear whether the results are specific to online CBT, with minimal therapist support, for child anxiety problems or whether they will generalize to online therapy alone, or indeed to clinic-based treatment. These remain questions for future research.

In summary, the present research demonstrated that a strong working alliance can be formed by parents and youths completing an online CBT intervention for anxiety. Indeed, we observed no significant differences in ratings of alliance in a clinic and online intervention for youths. Age significantly moderated the relationship between alliance, compliance, and outcome. For adolescents, in keeping with hypotheses, both working alliance and compliance positively and significantly predicted treatment outcome, and alliance significantly predicted compliance. However, compliance did not mediate the relationship between alliance and outcome. Younger children also tended to rate alliance very strongly, although alliance and compliance were not found to predict clinical outcomes. Taken together, the results highlight the importance of a strong client–therapist relationship in online therapy and support the use of online interventions for anxiety-disordered children and adolescents.

Conflicts of Interest

None declared.

References

1. Kendall PC, Flannery-Schroeder E, Panichelli-Mindel SM, Southam-Gerow M, Henin A, Warman M. Therapy for youths with anxiety disorders: a second randomized clinical trial. *J Consult Clin Psychol* 1997 Jun;65(3):366-380. [Medline: [9170760](https://pubmed.ncbi.nlm.nih.gov/9170760/)]

2. Wood JJ, Piacentini JC, Southam-Gerow M, Chu BC, Sigman M. Family cognitive behavioral therapy for child anxiety disorders. *J Am Acad Child Adolesc Psychiatry* 2006 Mar;45(3):314-321. [doi: [10.1097/01.chi.0000196425.88341.b0](https://doi.org/10.1097/01.chi.0000196425.88341.b0)] [Medline: [16540816](https://pubmed.ncbi.nlm.nih.gov/16540816/)]
3. Khanna MS, Kendall PC. Computer-assisted cognitive behavioral therapy for child anxiety: results of a randomized clinical trial. *J Consult Clin Psychol* 2010 Oct;78(5):737-745. [doi: [10.1037/a0019739](https://doi.org/10.1037/a0019739)] [Medline: [20873909](https://pubmed.ncbi.nlm.nih.gov/20873909/)]
4. March S, Spence SH, Donovan CL. The efficacy of an internet-based cognitive-behavioral therapy intervention for child anxiety disorders. *J Pediatr Psychol* 2009 Jun;34(5):474-487. [doi: [10.1093/jpepsy/jsn099](https://doi.org/10.1093/jpepsy/jsn099)] [Medline: [18794187](https://pubmed.ncbi.nlm.nih.gov/18794187/)]
5. Spence SH, Donovan CL, March S, Gamble A, Anderson RE, Prosser S, et al. A randomized controlled trial of online versus clinic-based CBT for adolescent anxiety. *J Consult Clin Psychol* 2011 Oct;79(5):629-642. [doi: [10.1037/a0024512](https://doi.org/10.1037/a0024512)] [Medline: [21744945](https://pubmed.ncbi.nlm.nih.gov/21744945/)]
6. Cartwright-Hatton S, Roberts C, Chitsabesan P, Fothergill C, Harrington R. Systematic review of the efficacy of cognitive behaviour therapies for childhood and adolescent anxiety disorders. *Br J Clin Psychol* 2004 Nov;43(Pt 4):421-436. [doi: [10.1348/0144665042388928](https://doi.org/10.1348/0144665042388928)] [Medline: [15530212](https://pubmed.ncbi.nlm.nih.gov/15530212/)]
7. Spence SH, Holmes J, Donovan CL. BRAVE for Teenagers-ONLINE: An Internet Based Program for Adolescents With Anxiety. Brisbane: University of Queensland; 2006.
8. Kazdin AE, Whitley M, Marciano PL. Child-therapist and parent-therapist alliance and therapeutic change in the treatment of children referred for oppositional, aggressive, and antisocial behavior. *J Child Psychol Psychiatry* 2006 May;47(5):436-445. [doi: [10.1111/j.1469-7610.2005.01475.x](https://doi.org/10.1111/j.1469-7610.2005.01475.x)] [Medline: [16671927](https://pubmed.ncbi.nlm.nih.gov/16671927/)]
9. Bordin EE. The Working Alliance: Basis for a General Theory of Psychotherapy. Washington, DC: American Psychological Association; 1975.
10. Baldwin SA, Wampold BE, Imel ZE. Untangling the alliance-outcome correlation: exploring the relative importance of therapist and patient variability in the alliance. *J Consult Clin Psychol* 2007 Dec;75(6):842-852. [doi: [10.1037/0022-006X.75.6.842](https://doi.org/10.1037/0022-006X.75.6.842)] [Medline: [18085902](https://pubmed.ncbi.nlm.nih.gov/18085902/)]
11. Horvath AO, Bedi RP. The alliance. In: Norcross JC, editor. *Psychotherapy Relationships That Work: Therapist Contributions and Responsiveness to Patients*. Oxford: Oxford University Press; 2002:37-69.
12. Kazdin AE, Marciano PL, Whitley MK. The therapeutic alliance in cognitive-behavioral treatment of children referred for oppositional, aggressive, and antisocial behavior. *J Consult Clin Psychol* 2005 Aug;73(4):726-730. [doi: [10.1037/0022-006X.73.4.726](https://doi.org/10.1037/0022-006X.73.4.726)] [Medline: [16173860](https://pubmed.ncbi.nlm.nih.gov/16173860/)]
13. Waddington L. The therapy relationship in cognitive therapy: a review. *Behav Cogn Psychother* 2002;30(2):179-192. [doi: [10.1017/s1352465802002059](https://doi.org/10.1017/s1352465802002059)]
14. McLeod BD. Relation of the alliance with outcomes in youth psychotherapy: a meta-analysis. *Clin Psychol Rev* 2011 Jun;31(4):603-616. [doi: [10.1016/j.cpr.2011.02.001](https://doi.org/10.1016/j.cpr.2011.02.001)] [Medline: [21482319](https://pubmed.ncbi.nlm.nih.gov/21482319/)]
15. 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](https://doi.org/10.1002/jclp.10263)] [Medline: [14981791](https://pubmed.ncbi.nlm.nih.gov/14981791/)]
16. Knaevelsrud C, Maercker A. Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: a randomized controlled clinical trial. *BMC Psychiatry* 2007;7:13 [FREE Full text] [doi: [10.1186/1471-244X-7-13](https://doi.org/10.1186/1471-244X-7-13)] [Medline: [17442125](https://pubmed.ncbi.nlm.nih.gov/17442125/)]
17. Cook JE, Doyle C. Working alliance in online therapy as compared to face-to-face therapy: preliminary results. *Cyberpsychol Behav* 2002 Apr;5(2):95-105. [Medline: [12025884](https://pubmed.ncbi.nlm.nih.gov/12025884/)]
18. Hawley KM, Garland AF. Working alliance in adolescent outpatient therapy: youth, parent and therapist reports and associations with therapy outcomes. *Child Youth Care Forum* 2008;37(2):59-74. [doi: [10.1007/s10566-008-9050-x](https://doi.org/10.1007/s10566-008-9050-x)]
19. Chu BC, Kendall PC. Positive association of child involvement and treatment outcome within a manual-based cognitive-behavioral treatment for children with anxiety. *J Consult Clin Psychol* 2004 Oct;72(5):821-829. [doi: [10.1037/0022-006X.72.5.821](https://doi.org/10.1037/0022-006X.72.5.821)] [Medline: [15482040](https://pubmed.ncbi.nlm.nih.gov/15482040/)]
20. Shirk SR, Karver M. Prediction of treatment outcome from relationship variables in child and adolescent therapy: a meta-analytic review. *J Consult Clin Psychol* 2003 Jun;71(3):452-464. [Medline: [12795570](https://pubmed.ncbi.nlm.nih.gov/12795570/)]
21. Karver M, Shirk S, Handelsman JB, Fields S, Crisp H, Gudmundsen G, et al. Relationship processes in youth psychotherapy: measuring alliance, alliance-building behaviors, and client involvement. *J Emot Behav Disord* 2008;16(1):15-28. [doi: [10.1177/1063426607312536](https://doi.org/10.1177/1063426607312536)]
22. Anonymous. Australian Bureau of Statistics. 2007 Dec 18. 8146.0: Household Use of Information Technology, Australia, 2007-08 URL: <http://www.abs.gov.au/ausstats/abs@.nsf/Previousproducts/8146.0Main%20Features42007-08?opendocument&tabname=Summary&prodno=8146.0&issue=2007-08&num=&view=> [accessed 2012-04-23]
23. Silverman WK, Albano AM. The Anxiety Disorders Interview Schedule for Children for DSM-IV. San Antonio, TX: Psychological Corporation; 1996.
24. Donovan CL, Holmes J, Spence SH. BRAVE for Teenagers: A Program for Adolescents With Anxiety. Brisbane: School of Psychology, University of Queensland; 2006.

25. Spence SH, Donovan CL, March S, Gamble A, Anderson R, Prosser SJ, et al. Online CBT in the treatment of child and adolescent anxiety disorders: issues in the development of BRAVE-ONLINE and two case illustrations. *Behav Cogn Psychother* 2008;36(spec iss 04):411-430. [doi: [10.1017/S135246580800444X](https://doi.org/10.1017/S135246580800444X)]
26. Tracey TJ, Kokotovic AM. Factor structure of the Working Alliance Inventory. *Psychol Assess* 1989;1(3):207-210. [doi: [10.1037/1040-3590.1.3.207](https://doi.org/10.1037/1040-3590.1.3.207)]
27. Feeley M, DeRubeis RJ, Gelfand LA. The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *J Consult Clin Psychol* 1999 Aug;67(4):578-582. [Medline: [10450629](https://pubmed.ncbi.nlm.nih.gov/10450629/)]
28. Hanson WE, Curry KT, Bandalos DL. Reliability generalization of working alliance inventory scale scores. *Educ Psychol Meas* 2002;62(4):659-673. [doi: [10.1177/0013164402062004008](https://doi.org/10.1177/0013164402062004008)]
29. Leibert T, Archer J. An exploratory study of client perceptions of Internet counseling and therapeutic alliance. *J Ment Health Counsel* 2006;28(1):69-84.
30. Busseri MA, Tyler JD. Interchangeability of the Working Alliance Inventory and Working Alliance Inventory, Short Form. *Psychol Assess* 2003 Jun;15(2):193-197. [Medline: [12847779](https://pubmed.ncbi.nlm.nih.gov/12847779/)]
31. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. *J Couns Psychol* 1989;36(2):223-233. [doi: [10.1037/0022-0167.36.2.223](https://doi.org/10.1037/0022-0167.36.2.223)]
32. Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, et al. A children's global assessment scale (CGAS). *Arch Gen Psychiatry* 1983 Nov;40(11):1228-1231. [Medline: [6639293](https://pubmed.ncbi.nlm.nih.gov/6639293/)]
33. Bird HR, Canino G, Rubio-Stipec M, Ribera JC. Further measures of the psychometric properties of the Children's Global Assessment Scale. *Arch Gen Psychiatry* 1987 Sep;44(9):821-824. [Medline: [3632256](https://pubmed.ncbi.nlm.nih.gov/3632256/)]
34. Dyrborg J, Larsen FW, Nielsen S, Byman J, Nielsen BB, Gautrè-Delay F. The Children's Global Assessment Scale (CGAS) and Global Assessment of Psychosocial Disability (GAPD) in clinical practice--substance and reliability as judged by intraclass correlations. *Eur Child Adolesc Psychiatry* 2000 Sep;9(3):195-201. [Medline: [11095042](https://pubmed.ncbi.nlm.nih.gov/11095042/)]
35. Rey JM, Starling J, Wever C, Dossetor DR, Plapp JM. Inter-rater reliability of global assessment of functioning in a clinical setting. *J Child Psychol Psychiatry* 1995 Jul;36(5):787-792. [Medline: [7559845](https://pubmed.ncbi.nlm.nih.gov/7559845/)]
36. Achenbach TM. Manual for the child behavior checklist/4-18 and 1991 profile. Burlington, VT: Department of Psychiatry, University of Vermont; 1991.
37. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. *J Abnorm Psychol* 1997 May;106(2):280-297. [Medline: [9131848](https://pubmed.ncbi.nlm.nih.gov/9131848/)]
38. Spence SH. A measure of anxiety symptoms among children. *Behav Res Ther* 1998 May;36(5):545-566. [Medline: [9648330](https://pubmed.ncbi.nlm.nih.gov/9648330/)]
39. Muris P, Schmidt H, Merckelbach H. Correlations among two self-report questionnaires for measuring DSM-defined anxiety disorder symptoms in children: the Screen for Child Anxiety Related Emotional Disorders and the Spence Children's Anxiety Scale. *Pers Individ Dif* 2000;28(2):333-346. [doi: [10.1016/S0191-8869\(99\)00102-6](https://doi.org/10.1016/S0191-8869(99)00102-6)]
40. Nauta MH, Scholing A, Rapee RM, Abbott M, Spence SH, Waters A. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. *Behav Res Ther* 2004 Jul;42(7):813-839. [doi: [10.1016/S0005-7967\(03\)00200-6](https://doi.org/10.1016/S0005-7967(03)00200-6)] [Medline: [15149901](https://pubmed.ncbi.nlm.nih.gov/15149901/)]
41. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986 Dec;51(6):1173-1182. [Medline: [3806354](https://pubmed.ncbi.nlm.nih.gov/3806354/)]
42. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr* 2009;76(4):408-420. [doi: [10.1080/03637750903310360](https://doi.org/10.1080/03637750903310360)]

Abbreviations

ADIS-C/P: Anxiety Disorders Interview Schedule for Children and Parents

CBT: cognitive behavior therapy

CGAS: Children's Global Assessment Scale

DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision

SCAS-C: Spence Children's Anxiety Scale-Child Version

SCAS-P: Spence Children's Anxiety Scale-Parent Version

WAI-S: Working Alliance Inventory-Short Form

Edited by F Thorndike, H Christensen, A Callear; submitted 29.05.11; peer-reviewed by M Khanna, A Parks; comments to author 11.07.11; revised version received 14.11.11; accepted 14.11.11; published 28.06.12.

Please cite as:

Anderson REE, Spence SH, Donovan CL, March S, Prosser S, Kenardy J

Working Alliance in Online Cognitive Behavior Therapy for Anxiety Disorders in Youth: Comparison With Clinic Delivery and its Role in Predicting Outcome

J Med Internet Res 2012;14(3):e88

URL: <http://www.jmir.org/2012/3/e88/>

doi: [10.2196/jmir.1848](https://doi.org/10.2196/jmir.1848)

PMID: [22789657](https://pubmed.ncbi.nlm.nih.gov/22789657/)

©Renee E E Anderson, Susan H Spence, Caroline L Donovan, Sonja March, Samantha Prosser, Justin Kenardy. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Viewpoint

Using Instructional Design Process to Improve Design and Development of Internet Interventions

Michelle M Hilgart¹, MEd; Lee M Ritterband¹, PhD; Frances P Thorndike¹, PhD; Mable B Kinzie², PhD

¹Behavioral Health and Technology, Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA, United States

²Curry School of Education, Department of Leadership, Foundation and Policy, University of Virginia, Charlottesville, VA, United States

Corresponding Author:

Lee M Ritterband, PhD

Behavioral Health and Technology

Department of Psychiatry and Neurobehavioral Sciences

University of Virginia

310 Old Ivy Way

Suite 102, PO Box 800223

Charlottesville, VA, 22903

United States

Phone: 1 (434) 924 5988

Fax: 1 (434) 244 7516

Email: LEER@virginia.edu

Abstract

Given the wide reach and extensive capabilities of the Internet, it is increasingly being used to deliver comprehensive behavioral and mental health intervention and prevention programs. Their goals are to change user behavior, reduce unwanted complications or symptoms, and improve health status and health-related quality of life. Internet interventions have been found efficacious in addressing a wide range of behavioral and mental health problems, including insomnia, nicotine dependence, obesity, diabetes, depression, and anxiety. Despite the existence of many Internet-based interventions, there is little research to inform their design and development. A model for behavior change in Internet interventions has been published to help guide future Internet intervention development and to help predict and explain behavior changes and symptom improvement outcomes through the use of Internet interventions. An argument is made for grounding the development of Internet interventions within a scientific framework. To that end, the model highlights a multitude of design-related components, areas, and elements, including user characteristics, environment, intervention content, level of intervention support, and targeted outcomes. However, more discussion is needed regarding how the design of the program should be developed to address these issues. While there is little research on the design and development of Internet interventions, there is a rich, related literature in the field of instructional design (ID) that can be used to inform Internet intervention development. ID models are prescriptive models that describe a set of activities involved in the planning, implementation, and evaluation of instructional programs. Using ID process models has been shown to increase the effectiveness of learning programs in a broad range of contexts. ID models specify a systematic method for assessing the needs of learners (intervention users) to determine the gaps between current knowledge and behaviors, and desired outcomes. Through the ID process, designers focus on the needs of learners, taking into account their prior knowledge; set measurable learning objectives or performance requirements; assess learners' achievement of the targeted outcomes; and employ cycles of continuous formative evaluation to ensure that the intervention meets the needs of all stakeholders. The ID process offers a proven methodology for the design of instructional programs and should be considered an integral part of the creation of Internet interventions. By providing a framework for the design and development of Internet interventions and by purposefully focusing on these aspects, as well as the underlying theories supporting these practices, both the theories and the interventions themselves can continue to be refined and improved. By using the behavior change model for Internet interventions along with the best research available to guide design practice and inform development, developers of Internet interventions will increase their ability to achieve desired outcomes.

(*J Med Internet Res* 2012;14(3):e89) doi:[10.2196/jmir.1890](https://doi.org/10.2196/jmir.1890)

KEYWORDS

Internet interventions; instructional design

Introduction

Emerging technologies have had far-reaching implications on global connectivity, including the availability and use of health information [1]. Given the accessibility and extensive capabilities of the Internet, it is increasingly being used to deliver comprehensive behavioral and mental health interventions. Their goals are to change user behavior, reduce unwanted complications or symptoms [2], and improve health status and health-related quality of life [3]. The Internet (and other technologies) offers the potential to provide efficient, interactive, tailored, and readily accessible health interventions [2].

Internet Interventions

Results of meta-analyses of Internet-delivered interventions show evidence of their efficacy for addressing a broad range of behavioral and mental health concerns [4-6]. Internet interventions have been found efficacious in addressing a variety of behavioral and mental health problems, such as insomnia, nicotine dependence, obesity, diabetes, depression, and anxiety [7-14]. Yet, despite the existence of many Internet-based interventions, the design of these interventions is widely variable, and there is little research to inform their design and development.

A model for behavior change in Internet interventions has been published [2] to help guide the development of Internet interventions and to help predict and explain behavior changes and symptom improvement through the use of these programs [2]. With this model, Ritterband and colleagues argue for grounding the development of Internet interventions within a scientific framework—a framework that explicitly identifies the importance of user characteristics, environment, intervention

content, level of intervention support, and targeted outcomes. More discussion, however, is needed regarding *how* the design of the program should be used to address these very issues.

Instructional Design Defined

While there is little research on the design and development of Internet interventions, there is a rich, related literature in the field of instructional design (ID). The term instructional design can be considered in three specific contexts. First, as a *science*, ID is concerned with how to help people learn more effectively. It includes research and theory about instructional, motivational, and behavioral learning strategies and the process models for designing and implementing instructional programs [15-18]. Second, as a *field of practice*, ID includes professional instructional designers working with teams of individuals to create detailed specifications for the development, design, implementation, evaluation, and maintenance of learning products [19]. These individuals make up part of the primary stakeholder groups for an intervention, together with other individuals who have an investment, or stake, in the success of the target population of learners (program users). Stakeholders include program developers, content experts, learners from the target population, and those affected by the program outcome. Finally, ID, as a *process*, (see Table 1) employs process models to guide the systematic development of instructional specifications drawing on learning, instructional, motivational, and behavioral theory to ensure the quality of instructional strategies.

Additionally, the ID process allows the translation of these theories into design principles that guide the development of the instructional product. This paper is focused on using ID as a systematic, reflective, and iterative process in the development of Internet interventions.

Table 1. The instructional design (ID) process: terms and definitions.

Instructional design	The systematic development of learning programs using theory to ensure the quality of instruction. It is the entire process of analysis of learning needs and goals, and the development of a delivery system to meet those needs.
ID process model	Prescriptive models that describe a systematic set of activities and steps involved in the planning, implementation, and evaluation of instructional programs.
Developmental research	The systematic study of designing, developing, and evaluating instructional programs, processes, and products that must meet the criteria of internal consistency and effectiveness.
Learners	The target population of an intervention: the individuals for whom the program or intervention is created.
Internet interventions	Internet-delivered, interactive, multimedia behavioral treatments often based on effective face-to-face interventions. Typically self-guided, highly structured, personalized, and tailored to the user to provide follow-up and feedback.
Stakeholders	Individuals who have a stake in the success of the target population of learners (program users). Stakeholders include program developers, content experts, learners from the target populations, and those affected by the program outcome.
Formative evaluation	The iterative process of tryout and revision of instruction and activities during development of the intervention before the actual implementation.
Needs assessment and analysis	Considers gaps between “what is” and “what should be” or “actual behaviors” versus “optimal behaviors.” Each gap is considered a need. A needs assessment or analysis is a process for determining how to close gaps. It involves identifying the required attitudes, behaviors, skills, and knowledge to meet needs.
Instructional goal	Formulated from the identified needs in the needs analysis. Instructional goals relate logically and persuasively to the documented performance gaps identified in the needs analysis.
Task analysis	Performed to identify the tasks required to reach the goals. This is an analysis of the content required for the desired instructional outcomes.
Learning objective	Written to specify exactly what the learner must do, know think, or feel as a result of completing the instruction. Objectives provide a framework for assessing and evaluating the extent to which learning is taking place.
Affective objective	Objectives that involve attitudes, emotions, and values.
Cognitive objective	Objectives and tasks related to information, knowledge, problem solving, and other intellectual aspects of learning.
Psychomotor objectives	Objectives that require the use of physical capabilities and activities, such as performing, manipulating, and constructing tasks.

Internet-Delivered Interventions and Instructional Design

During the past decade, educational researchers have focused on conducting design and development research to advance the practice of instructional development [15,20,21]. Studying the design and development process of an innovative instructional product can help developers better understand how to apply theoretical frameworks to the development process [22]. The same case can be made for studying the design and development of Internet-delivered interventions.

Developmental research in education seeks to create knowledge grounded in data systematically derived from practice [20]. By focusing specifically on the design and development process of creating Internet interventions, we can gain knowledge of the best practices and methods, and can develop more efficient models and frameworks for creating Internet interventions. Incorporating the ID process into the design and development of Internet interventions allows for continued testing and refinement of the theories that guide development and improvement in the resulting interventions. This is the essence of design-based research; theory dictates the design principles. Employing ID process during the design of Internet-delivered

interventions allows the programs to be evaluated for efficacy while also testing and refining the design principles that initially informed development of the intervention [23].

At their core, Internet interventions are innovative programs designed to teach skills, increase knowledge, and change behaviors, symptoms, or other targeted attributes. A majority of Internet interventions target specific behavioral, psychological, motivational, or health education outcomes, or a combination of these. Each intervention is implemented within its own context in which individuals from the target population (called learners or users) will use and (ideally) benefit from the program. It is important to clarify at the outset that the term learner or user is used to identify Internet intervention participants given their enrollment in a program designed to bring about change. These learners, however, are considered in different ways within the design process and the intervention itself (eg, patient or caregiver; child or parent; or children, teens, adults, or seniors). Although Internet-delivered intervention participants may not think of themselves as learners, they share the fundamental characteristic with learners of any intervention or involvement in a program with the aim of change, improvement, or advancement. This, together with the similarity of learning purposes for Internet-delivered health interventions

and other types of instructional interventions, makes the ID process an excellent fit for the design and development of Internet-delivered interventions.

Models of Instructional Design

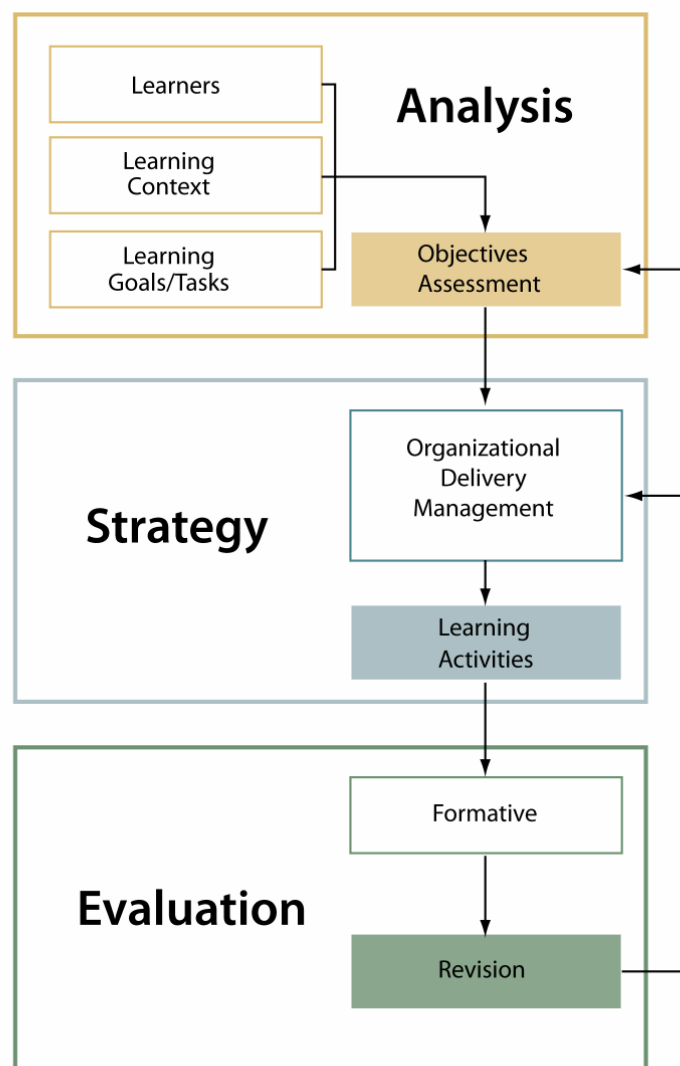
ID models can be used prescriptively to describe a systematic set of activities and steps involved in the planning, implementation, and evaluation of instructional programs. Using ID process models has been shown to increase the effectiveness of learning programs in a broad range of contexts, including both online and face-to-face formats [19,24-27]. ID models prescribe a systematic method for assessing the needs of learners to determine the gaps between current attitudes, behaviors, and knowledge and the desired outcomes [28]. The ID process guides designers to focus on the needs of learners, taking into account their prior knowledge [29]; set measurable learning objectives or performance requirements; design instructional strategies based on appropriate theory; assess users in a way that results in meaningful outcomes [30,31]; and use cycles of formative evaluation to ensure that the intervention meets the needs of all stakeholders.

We describe two ID models to convey the breadth in models from simple to elaborate, showing how ID models ultimately focus on similar activities. On one end of the complexity spectrum, Smith and Ragan [27] present what they call “a simple

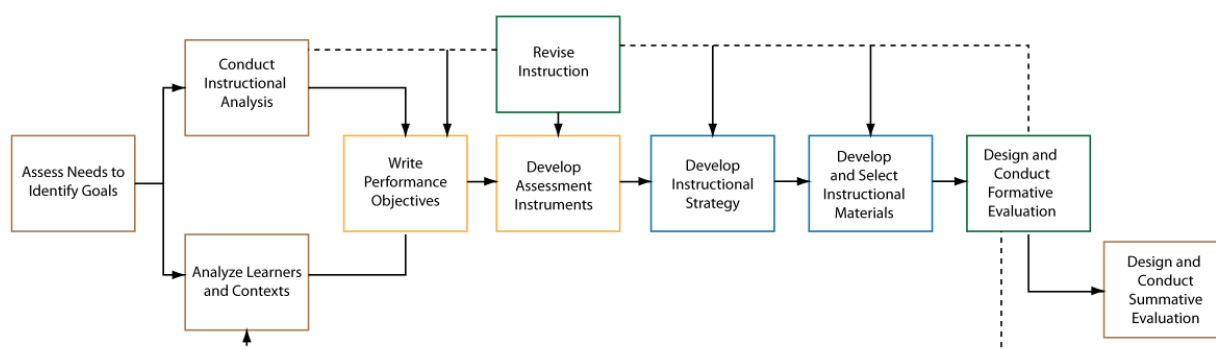
or common instructional design model” (see Figure 1) that focuses on three key activities or phases of the ID process: *analysis*, *strategy*, and *evaluation*. Analysis activities consist of assessing learners and learning contexts and developing learning goals. Strategy activities focus on design, organization, and delivery of instructional components. Evaluation activities focus on formative tryout of instruction to allow for revision before implementation.

Although the model is presented as linear, its authors point out that phases often happen concurrently, and considerations in one phase may (and often do) overlap with those in another phase. An important aspect of the ID process is that it is iterative. Formative evaluation begins during the first cycle, with “member checking” of the needs assessment, and continues throughout development. The results of the formative evaluation are used to make revisions to the intervention. The cycles repeat, with further evaluations guiding further revision.

In contrast to the simplicity of Smith and Ragan’s ID model [27], Dick and Carey’s Systems Approach Model for Designing Instruction (see Figure 2) presents a more complex ID model. This model displays more of the specific ID activities that take place [19], yet each of these activities can be mapped to the analysis, strategy, and evaluation components outlined in the Smith and Ragan model [27]. The key ID activities of analysis, strategy, and evaluation are described in detail below.

Figure 1. Smith and Ragan's instructional design process model.

Adapted from Smith and Ragan (2005)

Figure 2. The Dick and Carey Systems Approach Model for Designing Instruction.

Adapted from Dick, Carey, Carey (2001)

Analysis

In the analysis phase, the focus is on the targeted learner population (intervention users); the context and environment in which the learning (intervention) occur; identification of learning (intervention) goals and objectives; and the learning (intervention) tasks themselves. A needs assessment is performed to collect the information used in the analysis phase. Needs assessment and analysis considers the gaps between “what is” and “what should be.” A needs assessment is a process for determining how to close a learning or performance gap [32] and involves identifying the important needs and how best to meet them.

A particularly useful and relevant needs assessment approach, the discrepancy model of needs assessment, examines the gaps, or differences, between individuals who perform desired behaviors and those who do not [27]. For example, when a successful intervention requires learners to perform a behavior (take a medicine), use knowledge (recognize a skin problem), or show an attitude or belief (perceive one’s own risk), the differences between what ideal performers do, think, and feel are compared with what learners in the target population are actually doing, thinking, and feeling. Once gaps have been identified, the causes of the gaps can be studied and quantified. This information then shapes development of the learning or intervention tasks designed to achieve the intervention goals.

Internet intervention development is often based on effective face-to-face treatments where specific desired learning and performance outcomes have already been clearly identified and tested [33]. That is, Internet intervention developers frequently borrow from the goals identified in face-to-face treatments. Using the discrepancy approach to needs assessment, Internet intervention developers can focus on the causes of the gaps between what their target population is currently doing (knowing, thinking, or feeling) and what they should be doing to achieve the targeted outcome. Each gap discovered in the analysis is considered a need.

The discrepancy approach can also be used in instances where optimal behaviors have not been identified or are not known. In this case the optimal, or ideal, behaviors must be identified as part of the needs assessment process by collecting information from the target learners. This includes collecting information from those who have successfully achieved the desired outcome, as well as those who have not. Data are collected to inform developers about the level of awareness of the problem or condition of interest, the common symptoms, how risk is typically managed, and the level of adherence to preventive behaviors. As the data are analyzed, the attitudes, behaviors, skills, and knowledge of the successful individuals are identified and quantified as the optimal set of behaviors, skills, and knowledge. The differences between this ideal set and the unsuccessful set can then be identified as gaps or needs.

In the ID process, identified needs are then formulated into goals. It is critical that goals relate logically and persuasively to the documented performance gaps identified in the needs analysis. The importance of considering learner-based needs cannot be overstated. Programs must target the needs of the identified population in order to be successful [19,28].

Instructional and goal analysis (systematic methods for analyzing goals to identify the required knowledge, skills, behaviors, and attitudes to meet them) is used to categorize and prioritize goals, based on the kind of learning that will occur, into a series of specific measurable and observable objectives.

Objectives perform several critical functions in the design of instruction, including guiding designers toward the appropriate focus for instruction, and selecting activities and resources that facilitate effective learning [29]. Objectives also provide a framework for assessing and evaluating the extent to which learning is taking place and play an important role in guiding the learner by identifying the skills and knowledge to master [27,29].

During the analysis phase, the designer also considers the environment and context in which learning will take place. The environmental approach to analysis is based on three environmental domains: physical, social, and institutional [34]. Physical concerns are those related to the physical environment in which the intervention will be used. Social concerns refer to the learners and their social connections and networks, including those that will influence the learning experience. Institutional considerations affect any institutional goals held by the sponsoring organization and help define the dissemination and use of the program. Each of these domains informs decisions that affect a learner’s ability and willingness to access and use the program. By working closely with members of the targeted population at the early stages of design and development, designers can fully consider the domain factors most relevant for creation of programs that reflect and address the needs of learners, as well as aligning with the domains in which the programs will be used.

Task analysis, the next step in the ID process, is conducted to identify the actual tasks required to reach the goals identified in the needs analysis. This second analysis considers the content required for the desired instructional outcome [29]. Although there are many ways to analyze tasks, most ID models provide a scheme for classifying information into discrete categories [15]. The objectives and tasks typically fall into one of three domains, or categories: cognitive, psychomotor, and affective [29]. The cognitive domain includes objectives and tasks related to recall of information, development of conceptual knowledge, application of knowledge to problem solving, and other intellectual aspects of learning. The psychomotor domain includes skills that require the use of physical capabilities and activities, such as performing, manipulating, and constructing tasks. The affective domain includes objectives targeting attitudes, emotions, and values. Thoughtful consideration of the learning experience, including the cognitive, psychomotor, and affective factors involved in that experience, will increase the likelihood that learners can successfully develop the desired knowledge, skills, behaviors, or attitudes.

In summary, the analysis phase of the ID process involves two sets of analyses. The first analysis identifies the learners (intervention users), the learning (intervention) objectives, and the environment in which the learning (intervention) occurs. The second analysis considers the content, type of tasks, and learning experiences required to meet learning objectives.

Strategy

To guide creation of the instructional activities, the strategy phase of the ID process is informed by and draws upon tested theories. Theories are drawn from a range of fields including education and learning psychology, behavior change, and motivation [16-18,35-38]. Theories that have been widely applied include cognitive learning [18,35,39,40], information processing [41], and multimedia learning [42]. The structure and type of learning required by each objective influences which learning theories are most applicable. The selection of instructional strategies is also clearly influenced by the analysis of the content; that is, the determination of domain (ie, whether the target is a behavior, skill, knowledge, or attitude), as well as the analysis of the tasks that make up the desired performance. For example, if the intervention aim is to teach the users a rule (eg, to get out of bed if they have not fallen asleep in 20 minutes as part of an intervention for adults with insomnia), the developer first recognizes that this knowledge acquisition falls in the cognitive domain, and then considers the cognitive skills that lead up to and support the rule application (eg, conceptual understandings, which are in turn supported by information recall). Developers would then turn to cognitive theories to help inform their development of the learning activities designed to facilitate learning to apply this rule. Different types of learning tasks require substantially different levels of cognitive effort and different kinds of learning conditions [24,27,29].

In this strategy phase of the ID process, the focus is on designing the learning activities that will best serve the specific set of learners for whom the program is being developed. Learning activities refer to learning experiences that involve informational content and designed experiences in which learners act on content in specific ways. Consideration is given to each specific objective and how best to actively engage learners with the learning experiences in order to obtain the desired result. The crafting of instructional strategies is considered the most crucial step in the ID process; it is the step that can contribute the most to making instruction successful [29].

When selecting instructional strategies, designers also need to consider and select the media and methods best suited to deliver the desired experience. Smith and Ragan's model [27] (see Figure 1) highlights three key categories in the strategy phase: organization, delivery, and management. Organizational strategies focus on how instruction will be sequenced and presented. Delivery strategies are concerned with the instructional media that will be used and how learners are grouped. For example, learners may be grouped by level of prior knowledge, attitudes on the topic, skill abilities, motivation level, or presence of specific symptoms. In selecting the appropriate media elements used for the learning activities (eg, text, audio, graphics, and animation), developers should evaluate the motivational appeal and ability of each element to support learners in recalling prior knowledge, providing new learning stimuli, activating responses, providing feedback, and encouraging practice and transfer [16,19,43]. Finally, management strategies focus on the scheduling and implementing of instruction.

In sum, strategy activities are critical to the ID process. They include the design and development of the actual learning and behavior change activities that will be used to help achieve the objectives and offer designers the opportunity to draw on strategies that have previously been shown to be effective within other specific contexts. The strategy phase also focuses on the organization or sequencing of learning activities, and the media and delivery methods used to engage learners with the materials.

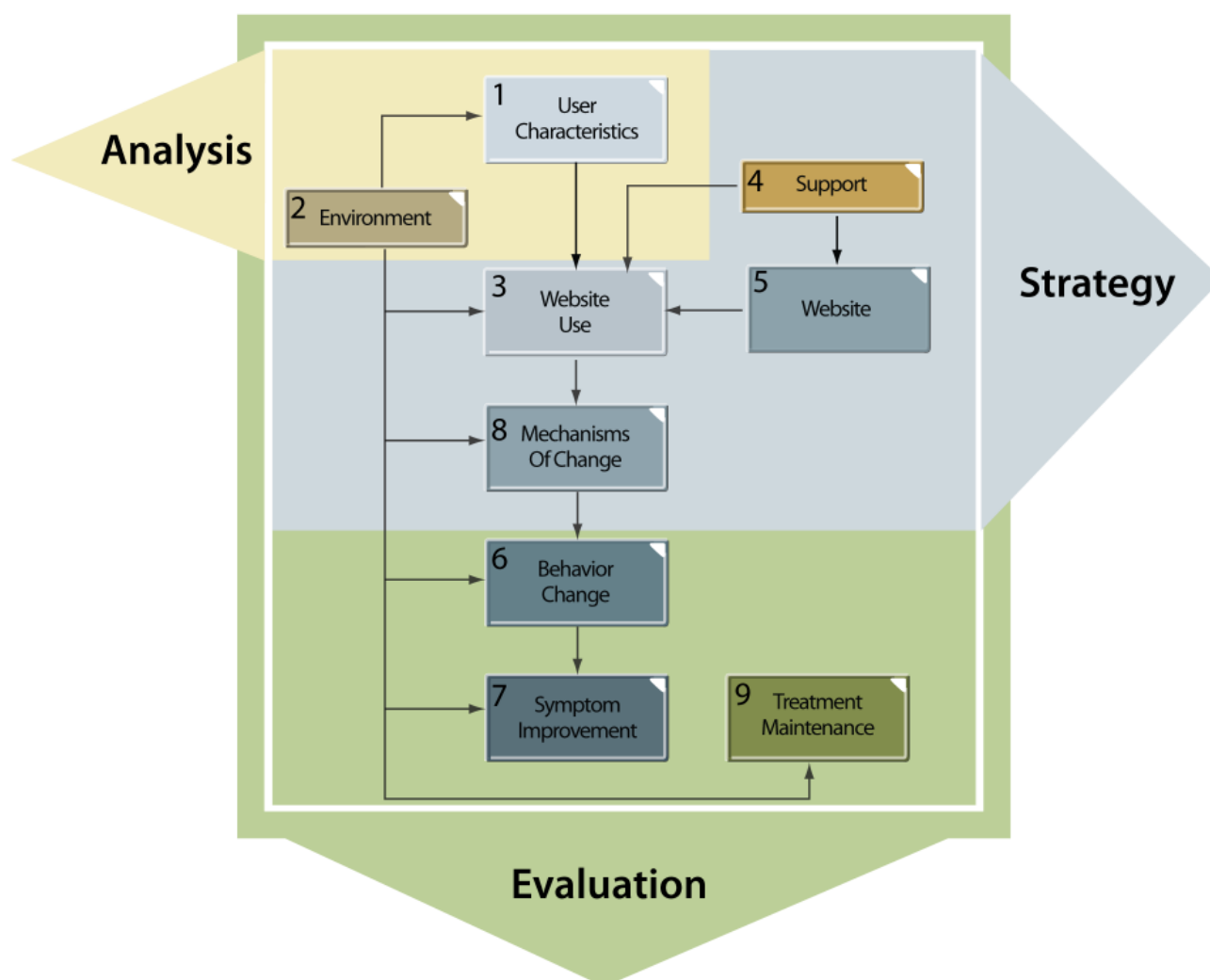
Evaluation

Formative evaluation refers to the iterative process of tryout and revision of activities or content during development before the actual implementation [44]. The evaluation phase tests the assumptions made in the analysis and strategy phases. The purpose of formative evaluation is to ensure that the goals of the instruction are being achieved and to revise the program as needed before implementation. Formative evaluation requires a plan for determining the extent to which learning is taking place. It involves trying out learning activities with members of the learning group. Evaluation instruments are used to assess the learner's mastery of the objectives.

Formative evaluation ideally takes place at all stages of the ID process. In fact, Dick et al incorporate a formative evaluation element into each stage of their ID process model [19] where draft versions of the instruction are examined and then revised as needed. Formative evaluation during the earliest stages of the ID process (even needs assessments can be member checked and refined for fit by stakeholders) can help determine whether the learning goals and objectives have been correctly identified, and whether assumptions made about learners and learner characteristics hold true. This helps prevent valuable time and resources from being wasted on components that are not effective. Using an ongoing formative evaluation approach, and revising the program based on findings, allows developers to identify weaknesses that can be corrected before full-scale implementation [19,24,29,44]. It also increases the probability that the program, when implemented, will produce the desired learning and performance outcomes. Failure to conduct formative evaluation throughout all stages of the ID process misses opportunities for identification of problem areas and increases the possibility that learners will not achieve the intended goals despite considerable investment of resources.

Integrating ID Process Into the Design and Development of Internet Interventions

Ritterband and colleagues [2] described a behavior change model for Internet interventions (see Figure 3 for a high-level representation of the Internet intervention model juxtaposed with the ID model) that consists of nine nonlinear steps: (1) the *user*, influenced by (2) *environmental factors*, affects (3) *website use* and adherence, which is influenced by (4) *support* and (5) *website characteristics*. *Website use* leads to (6) *behavior change* and (7) *symptom improvement* through various (8) *mechanisms of change*. The improvements are sustained via (9) *treatment maintenance* or relapse prevention. Each section of this behavior change model can be clearly connected to the ID process.

Figure 3. Instructional design process model for behavior change in Internet interventions.

User Characteristics and Environmental Factors

Ritterband and colleagues [2] identify seven areas of user characteristics that are congruent to the analysis phase of the ID model, including disease, demographics, traits, cognitive factors, beliefs and attitudes, physiological factors, and skills. Using an ID approach focuses the needs assessment on discrepancies between the user's *desired* behaviors, skills, knowledge, or attitudes and their *current* behaviors, skills, knowledge, or attitudes. Identifying these discrepancies accomplishes two important functions: (1) it attempts to quantify the current state of affairs with regard to the target population so that progress toward meeting goals can be accurately measured [24], and (2) it allows very specific learning and performance objectives to be crafted for the intervention that are based on the desired outcomes.

Environmental factors are also considered in the analysis phase of the ID model. As in Ritterband and colleague's model [2], environment from the ID perspective is composed of multiple

influences, and the focus on environment is holistic. Assessment methods are used to construct an environmental snapshot of how the program will be used by the learner from the physical, social, and institutional contexts of the learning environment. When using the ID model, designers can identify potential program supports and barriers and design the program accordingly. For example, through a needs assessment, designers of an Internet intervention targeting users with negative attitudes toward school are alerted that labeling recommended intervention activities as homework or assignments may present barriers for these users. The ID model also allows designers to leverage environmental factors that influence user characteristics. By systematically identifying the environmental factors, along with the characteristics of the user population, program designers can choose the most relevant motivation, learning, and behavior change theories for their set of learners and learning objectives based on the research literature. This, in turn, will affect the type of instructional strategies that are planned in the strategy phase.

Website Use, Website Characteristics, Support, and Mechanisms of Change

The website use, website characteristics, support, and mechanism of change components of the Internet intervention model all map to the strategy phase of the ID process model. All phases of the ID process are interrelated, so the findings and assumptions made in the analysis phase greatly affect the decisions made in the strategy phase. The emphasis in the strategy phase is on crafting the instructional strategies used to reach the desired goals. This focus underscores the importance of design-based research. Using a design-study approach, developers of Internet interventions tap theory for instructional strategy design guidelines that are then tested along with the instructional components. Thus, knowledge is created in the refinement of theories of both design or development and instruction in the context of Internet-delivered interventions.

There are several specific challenges in considering instructional strategies in the context of Internet interventions. Opportunities to leverage technologies to build activities that support and promote learning are often not fully considered in favor of less expensive and more quickly produced solutions. For example, text-based delivery of learning is recognized as an important method, yet, before a text-focused solution is selected as most appropriate, it must be considered in relation to the target population of learners, their prior experiences and attitudes, and their reading levels. When text is used, learners can be grouped based on reading level or supported by providing narration so those with lower-level reading skill have the choice of listening instead of reading. Video clips, animation, audio segments, and images can all be integrated with text to support meeting learning objectives.

Multimedia learning and using specific interactions to affect motivation, skill building, and behavioral change demand instructional strategies that move beyond text on a screen. These methods can be particularly appropriate when an intervention includes psychomotor skill building. Modeling, through the use of videos or animation of specific skills, using a variety of models and conditions can be more effective than reading about the activity [18,42]. Additionally, modeling typical beginner mistakes, while providing specific feedback on why problems occur and ways to correct or prevent them, can help learners gain mastery [30].

Motivation, or the action of an individual to select and sustain a behavior, is another theory area from which designers of instructional strategies can find significant guidance. Motivation theories help focus attention and support affirmative answers to questions such as “Can I do this task?,” “Do I want to do this task?,” and “Will I continue to do this task?” Self-efficacy [45], expectancy-value [17], self-determination [46,47], and self-regulated learning theories [47] can all provide designers with practical and applicable programmatic supports that can help learners meet learning objectives and help sustain adherence to intervention use.

The strategy phase also focuses on organizing and sequencing learning activities in ways that will best help users meet identified program goals. Strategy activities include selecting media and methods for delivery of the instruction, and managing

and supporting the implementation of the instructional strategies. In each of these areas, the primary concern is how best to engage the learner to reach agreed-on goals.

Behavior Change, Symptom Improvement, and Treatment Maintenance

The formative evaluation component of the ID model is the evaluation of program components completed while the intervention is being formed. This type of evaluation provides evidence for how well designers have reached their goals and allows modification of the program before it is fully implemented. Behavior change, symptom improvement, and treatment maintenance or relapse prevention components of the Internet intervention model align with the evaluation phase of the ID process. During formative evaluation activities, results of the needs assessment, instructional goals, objectives, and strategies are evaluated with content experts and members of the target population to explore whether the assumptions made, the strategies employed, and the learning activities developed actually result in the desired outcomes.

To determine the extent to which goals are being met, designers develop and implement a plan for assessing outcomes from their intervention. Evaluation of instruction typically considers outcomes at four levels: (1) learner reactions, (2) learning achievement, (3) transfer of learning, and (4) organizational results [26]. Formative evaluation focuses on the first two levels [24], while summative evaluation (program evaluation that occurs after implementation) may focus on all four levels. The first level, learner (user) reactions, refers to the attitudes and preferences of the learners toward the learning intervention. This is the extent to which learners like or dislike the learning activities, or find the activities to be satisfactory, effective, and useful. The second level, learning (intervention) achievement, refers to how well learners perform on objective measures of learning. This is often tested with pre and post tests to determine the extent to which learners have mastered goals via change in attitudes, knowledge, skills, or behavioral intentions.

During the revision stage of formative evaluation, the data that have been collected are analyzed and used to operationalize a set of revisions to the intervention. The ID process includes setting standards and criteria to guide revisions. This includes criteria for examining the data that have been collected; criteria for organizing and summarizing the data; and criteria for prioritizing which sources of data are most relevant to the revision efforts [44]. Prioritizing and deciding how to implement revisions is typically the most challenging for designers. It can be relatively straightforward to identify a problem area but less clear what should be the appropriate revision or refinement of the instruction. For example, an evaluation of learners on how well they like an intervention and find it satisfactory and useful (level 1) returns negative results for one set of learners but positive results for other sets. Designers need to think about how the learners are being grouped and whether the set of learners with the unsatisfactory experience share characteristics that can help inform program revisions (eg, gender differences, differences in prior knowledge, differences in symptoms, or age-related differences). Designers also need to consider the selected theoretical underpinning and whether it is appropriate

for the subset of learners with negative outcomes. Designers should investigate whether additional needs within a subset of the user population can be identified and supported by adding appropriate objectives, content, and instructional strategies, thus making the intervention satisfactory, useful, and relevant to all user groups. It is often necessary to reconsider the previous stages of analysis and strategy to determine how best to revise the instruction [19].

By conducting preliminary testing of the Internet intervention with users from the target population, designers will be able to measure the extent to which they have reached their objectives and to further refine goals or strategies to best reach the desired outcomes. All elements of the Internet intervention model are incorporated because formative evaluation is conducted at each phase of the ID process to confirm the assumptions made in that phase. Another important aspect to note is that, because the ID process is highly interconnected, all elements of the model overlap considerably. For example, while this mapping shows behavior change, symptom improvement, and treatment maintenance as mapping only to the evaluation phase of activities of the ID process, there is also a clear connection to both analysis and strategy activities.

Application

Here are several examples to help convey the process of incorporating ID into the development of Internet interventions. These examples are broken down into the analysis, strategy and evaluation phases.

Analysis Phase

Instructional curriculum mapping is an ID method that uses flowcharting to illustrate instructional relationships within a program [48]. An example of the use of instructional curriculum mapping flowcharts is seen in the development of iSHIFTup (Internet Skin Health Intervention For Targeted Ulcer Prevention), an Internet-delivered intervention designed for adults with spinal cord injury to prevent serious pressure ulcers (funded by the Commonwealth Neurotrauma Initiative, Virginia, USA). Within iSHIFTup, instructional curriculum mapping flowcharts were used to show how objectives map to instructional sequences. Figure 4 shows a visual representation of the relationship between an objective for learners in iSHIFTup “to identify personal risk factors for pressure ulcers” and the skills and attitudes required to meet the objective. In this example, the objective skill is within the cognitive domain. The desired outcome is that learners be able to identify their own risk factors. To do this, learners first acquire the skills to

describe general risk factors, then compare their own behaviors with these factors and classify their risk factors. Once learners have mastered these sequences, they are able to identify personal risk factors for pressure ulcers.

Figure 5 (a small section of a larger, fully developed iSHIFTup flowchart) graphically shows learners’ movement through the intervention from their first contact with the program to completion of all required intervention components. Flowcharting allows communication of the complex timing of events used in Internet interventions. Using these systematic methods allows visualization of the entire intervention and facilitates discussion with all stakeholders (including members of the target population of learners) to help reach consensus on the process. Using flowcharts to envision the whole system as it is being designed supports coordination between designers, developers, and those who will implement the program [19,27]. It also allows for a common language and general procedure among the stakeholders. These flowcharts, and the written plans that accompany them, are results of the ID method. They aid the development team in the process of iterative review and revision of work in a coordinated and systematic manner.

Another example of ID process in the analysis phase is the inclusion of learners from the target population in the design of the intervention. Including members of the group who will use the program in its design is critical to the success of the intervention. For example, in the design of iSHIFTup, 10 individuals (8 with spinal cord injury and 2 caregivers of patients with spinal cord injury) partnered with the development team. In addition to the 10 target user members, the development team consisted of content experts in the area of Internet-delivered interventions, pressure ulcers, and spinal cord injury medicine (clinical psychologist, rehabilitation physician, wound specialist, physical therapist, and occupational therapist), an instructional designer, a graphic artist, and a Web programmer. Collaboratively, the team identified program goals, content, and functional requirements for the intervention. Individuals with spinal cord injury shared their real-life stories about the challenges of keeping their skin healthy and maintaining preventive behaviors. They shared critical information about living with pressure ulcers and identified which coping strategies had (and had not) worked for them. Members of the user population also reviewed intervention content and instructional activities throughout the design and development of the program. They gave feedback that was incorporated into program revisions, and later reviewed the revised content. This collaboration serves to ensure relevance and acceptability to the targeted learners (users) of the intervention.

Figure 4. Example of instructional curriculum mapping. Program objective: core level. IS = intellectual skill; VI = verbal information.

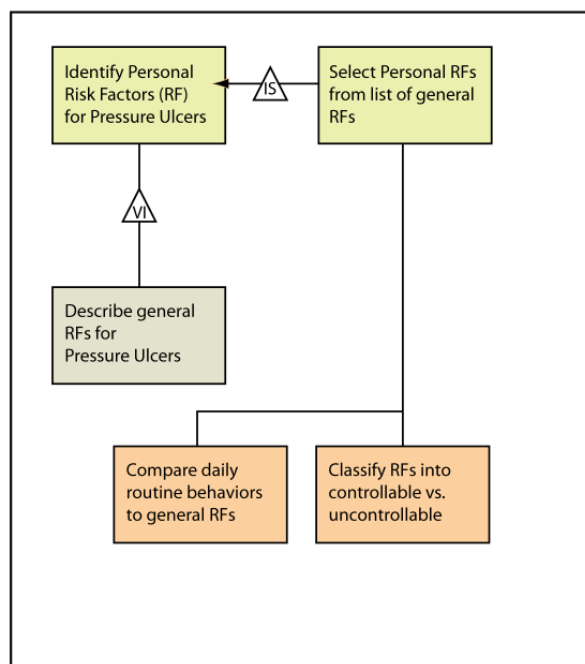
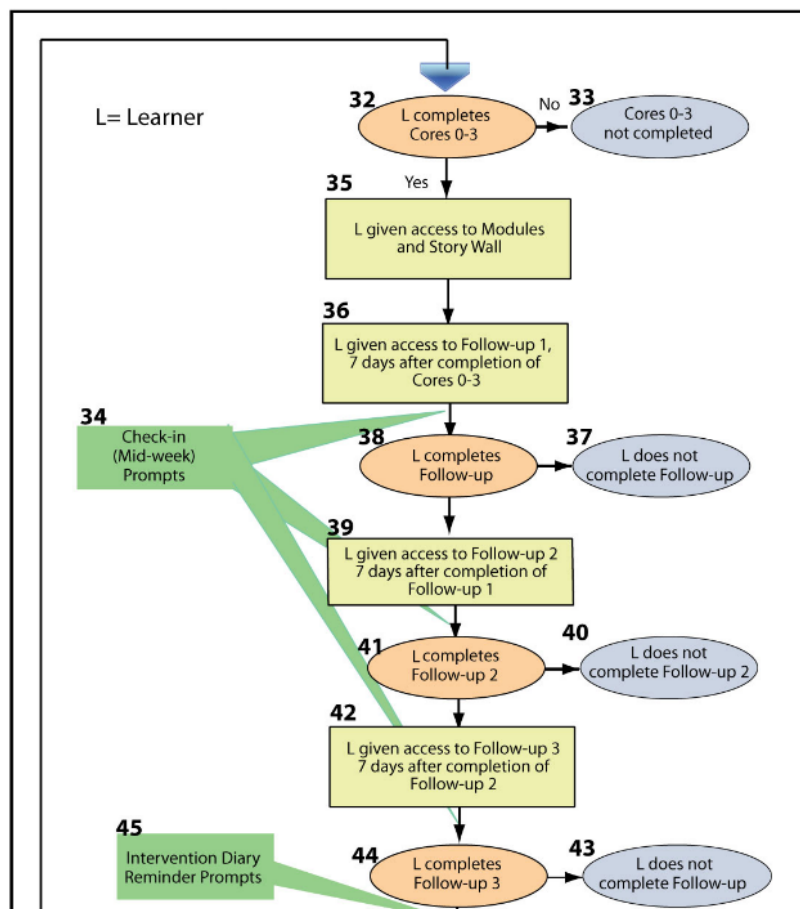


Figure 5. Example of instructional curriculum mapping. Program overview: sequence.



Strategy Phase

A comprehensive set of theory-driven instructional strategies has been recommended for health education [49]. A condensed version of Gagne's events of instruction [43] is used as a starting point, or framework, on which the health belief model [50], social cognitive theory [18], and diffusion theory [51] are drawn as key health behavior change theories to inform instructional strategy development. Specific instructional practices based on these theories are advanced for each of Gagne's instructional events. For example, in the recommended strategy for Gagne's event *provide learner guidance*, social cognitive theory suggests using trustworthy, knowledgeable modeling to demonstrate desired behaviors and social modeling to develop self-protective skills. For example, a series of photographs or illustrations of a trusted, competent person performing preventive or protective behaviors could be used. Diffusion theory suggests using trustworthy and knowledgeable opinion leaders from the target population to speed the diffusion process. An example of an instructional strategy informed by both of these theories would be a video of a recognized opinion leader (eg, well-known athlete, musician, or actor), who is identified with the target population, performing desired behaviors, such as checking blood glucose levels before driving.

Another example of using ID process during the strategy phase can be seen in SHUTi (Sleep Healthy Using The Internet), an Internet intervention for adults with insomnia [7]. One of the learning goals in this intervention is that users "recognize the relationship between *Time in Bed* and *Sleep Efficiency*." This relationship has been identified as a conceptual understanding required to apply the behavioral rules of sleep restriction. Sleep efficiency is defined as the mathematical calculation of *total sleep time* divided by *time in bed*, multiplied by 100. In SHUTi, learners are cued (using highlighting and color) to move sliders to set and subsequently increase and decrease their total sleep time and time in bed to graphically see the relationship between the two. For example, as time in bed (while not asleep)

increases, sleep efficiency decreases; and as time in bed approaches total sleep time, sleep efficiency increases (Figure 6).

This interactive, user-controlled activity was selected as an ideal way to engage learners, based on their characteristics (above-average education, high motivation, and comfort with technology), the content being introduced (cognitive domain, making connections, and intellectual skill), and type of learning goal (to recognize the relationship between time in bed and sleep efficiency and apply it to their own situation). Several learning theories were applied in crafting this solution. Guided discovery e-learning architecture [52] was selected in which knowledge construction is the learning goal and high interactivity is used to guide learners to specific goals such as making connections and identifying relationships. Multimedia learning theory suggests that people learn better when graphics are used to show relationships [53]. Here, the quantitative relationship between time in bed, total sleep time, and sleep efficiency is shown graphically. By adding interactivity in the form of slider bars, it becomes a transformational graphic that depicts changes over time [42]. The theory of planned behavior [54] is used to encourage learners to consider their own behaviors (time in bed) and intentions.

These examples demonstrate the ID process of drawing on learning theory and applying it to the specific learning context in which the intervention is being used. Design-based research, which blends empirical instructional research with the theory-driven design of learning environments [55], is an important methodology for discovering which learning and behavior theories work in Internet interventions and under what conditions. This can lead to sharable theories that help communicate relevant implications to developers of Internet interventions [23,55]. Using this type of design-based research approach focuses on methods that document processes of enactment to outcomes of interest.

Figure 6. SHUTi (Sleep Healthy Using The Internet) sleep efficiency interaction screen.

Evaluation

The final example focuses on formative evaluation and revision of instruction as part of the evaluation phase of ID. This example is an Internet-delivered intervention designed for pediatric encopresis (UCanPoopToo) [8,33]. In this program, learners (parent and child dyads) using the intervention took part in formative evaluation activities. Analysis of the findings according to agreed-on criteria revealed that parents using the program could be further supported by including a self-assessment to determine whether their children had mastered each unit of instruction. Once this gap had been recognized, additional learning goals were identified to support parents in assessing their children's mastery. Self-regulated learning theory

[47] was used to inform the design of self-assessments of the child's content mastery and to allow parents to reflect on and adapt their children's learning processes toward the learning goals.

In the revision phase of the formative evaluation, the design team used core objectives as a starting point for developing the self-assessments. The result (see Figure 7) is a self-assessment at the end of each unit of instruction called *Now I Can*. Each Now I Can screen describes specifically what the child should be able to do at the end of the instruction. In current versions of the program, parents complete Now I Cans with their child to assess whether their child has mastered the skills in that core of instruction.

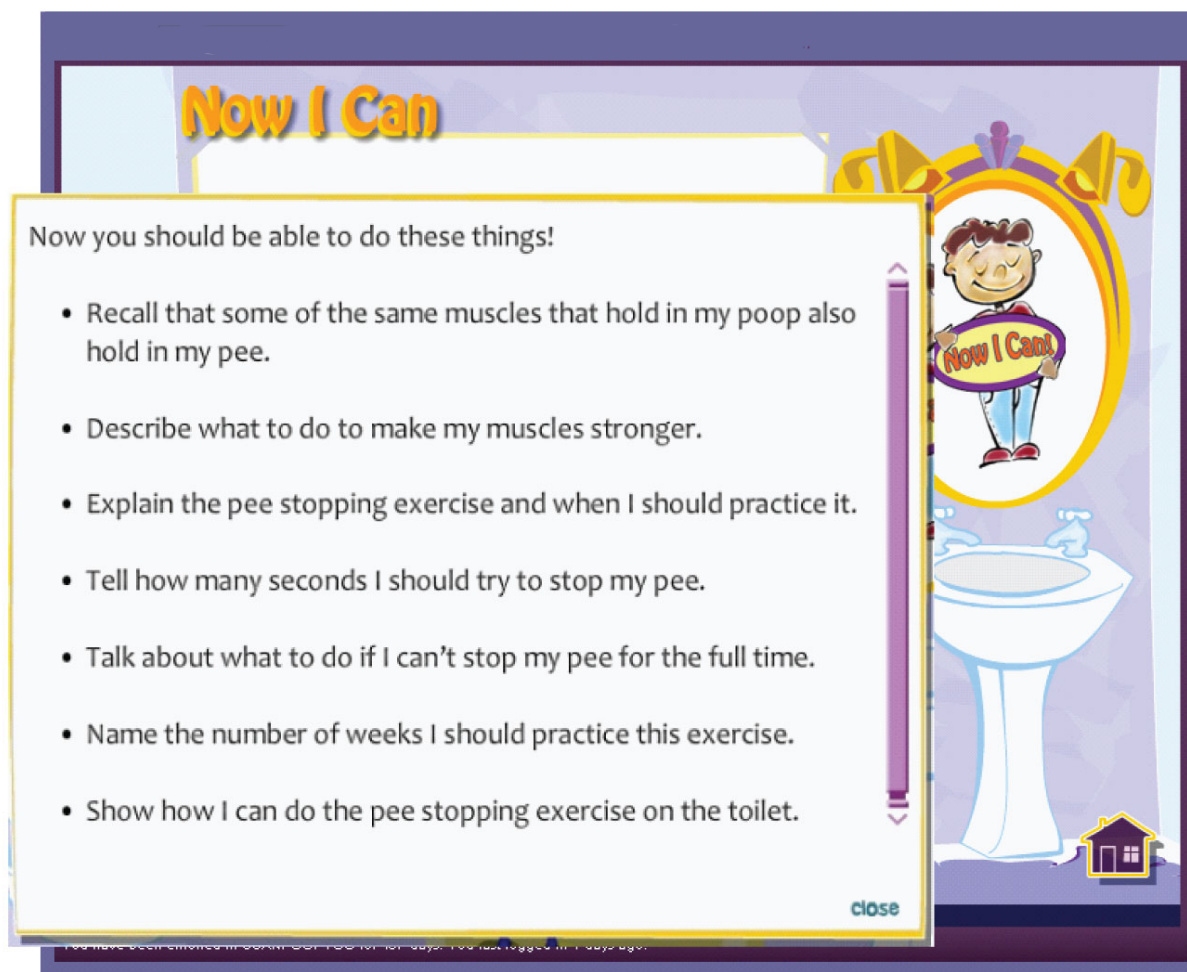
Figure 7. Now I Can screen from UCanPoopToo.

Figure 8. Internet intervention instructional design check list.

ANALYSIS	Yes	No	N/A
Were learner/user characteristics considered?			
Demographics			
Disease or health issues			
Traits			
Cognitive factors (prior knowledge, reading level)			
Physiological factors			
Beliefs and attitudes			
Motivation toward topic/tasks			
Was learning context considered?			
Physical: Conditions in which intervention is used			
Social: Relationships that impact intervention use			
Institutional: Leadership support for intervention			
Was a discrepancy needs approach considered?			
Optimal			
Required attitudes			
Required knowledge			
Required skills			
Required behaviors			
Gaps			
Needs (gaps between optimals and actuals) described and quantified			
Needs prioritized and written as goals			
Were goals used to create specific objectives?			
Determine types of learning to meet goals			
Write specific learning objectives (should be measurable and observable)			
Identify content to meet learning objectives			
Develop assessment of objectives			
STRATEGY			
Were delivery, organization, and management considered?			
Identify strategies based on types of learning			
Select media and methods best suited to objectives and learners			
Determine how instructional activities will be sequenced and presented			
Determine how learners will be grouped (prior knowledge, skill level, etc.)			
Determine how intervention will be implemented and sustained			
Were appropriate theories used to inform strategies?			
EVALUATION			
Was a plan made to determine whether goals are being met?			
Learner reaction to the intervention (satisfaction): Level 1			
Learning achievement from the intervention (pre/post): Level 2			
Transfer of learning (use in real-life, change in behaviors): Level 3			
Was collected data used to revise the intervention?			
Set standards and criteria to guide revisions			
Prioritize and implement revisions			
Revisions to intervention should also be evaluated			

Conclusion

Developers of Internet interventions often struggle with the question of whether adding a particular feature to an intervention, such as a game, a simulation, or animation, is a good or bad idea. This, however, is not the salient question to consider. Instead, given the concepts of the ID process set forth here, designers of Internet-delivered interventions are encouraged to take a learner-centered, needs-based approach and to consider how *all* technology features (eg, text, graphics, interactivity, video, and games) can be used in ways to best meet the needs of learners. It is not a simple question of whether to include a feature, but is instead a broader, more complex question of what theory-based learning activities best support a specific set of learners given their own characteristics and learning environment to meet an identified set of measurable objectives.

The ID process offers a proven methodology for the design of instructional programs and should be considered an integral part of the creation of Internet interventions. To support researchers, an Internet intervention ID check list has been created and included to use in developing new Internet-delivered interventions (see [Figure 8](#)). By providing a framework for the design and development of Internet-delivered interventions and by purposefully focusing on the design, development, and the underlying theories supporting these practices, both the theories and the interventions themselves will continue to be refined and improved. By using the behavior change model for Internet interventions along with the best research available to guide design practice and inform development, developers of Internet-delivered interventions will increase their ability to help users achieve the desired outcomes.

Conflicts of Interest

Drs. Ritterband and Thorndike are equity owners of BeHealth Solutions, Inc, which is developing products related to the research described in this paper. The terms of this arrangement have been reviewed and approved by the University of Virginia, in accordance with their respective conflict of interest policies.

References

1. Fox S. Online health search 2006. Washington, DC: Pew Internet & American Life Project; 2006 Oct 29. URL: http://www.pewinternet.org/~media/Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf [accessed 2012-05-01] [[WebCite Cache ID 67L69VTfk](#)]
2. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. *Ann Behav Med* 2009 Aug;38(1):18-27. [doi: [10.1007/s12160-009-9133-4](https://doi.org/10.1007/s12160-009-9133-4)] [Medline: [19802647](#)]
3. 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](https://doi.org/10.1146/annurev.publhealth.031308.100235)] [Medline: [19296777](#)]
4. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. *J Technol Hum Serv* 2008;26(2):109-160. [doi: [10.1080/15228830802094429](https://doi.org/10.1080/15228830802094429)]
5. Spek V, Cuijpers P, Nyklíček 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](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](#)]
6. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. *J Med Internet Res* 2004 Nov 10;6(4):e40 [[FREE Full text](#)] [doi: [10.2196/jmir.6.4.e40](https://doi.org/10.2196/jmir.6.4.e40)] [Medline: [15631964](#)]
7. Ritterband LM, Thorndike FP, Gonder-Frederick LA, Magee JC, Bailey ET, Saylor DK, et al. Efficacy of an Internet-based behavioral intervention for adults with insomnia. *Arch Gen Psychiatry* 2009 Jul;66(7):692-698 [[FREE Full text](#)] [doi: [10.1001/archgenpsychiatry.2009.66](https://doi.org/10.1001/archgenpsychiatry.2009.66)] [Medline: [19581560](#)]
8. Ritterband LM, Ardan K, Thorndike FP, Magee JC, Saylor DK, Cox DJ, et al. Real world use of an Internet intervention for pediatric encopresis. *J Med Internet Res* 2008;10(2):e16 [[FREE Full text](#)] [doi: [10.2196/jmir.1081](https://doi.org/10.2196/jmir.1081)] [Medline: [18653440](#)]
9. Cox DJ, Ritterband LM, Gonder-Frederick L, Patel K, Clarke W. Blood glucose awareness training delivered over the internet. 2006 Presented at: American Diabetes Association 66th Scientific Sessions; June 9-13, 2006; Washington, DC, USA.
10. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. *JAMA* 2003 Apr 9;289(14):1833-1836 [[FREE Full text](#)] [doi: [10.1001/jama.289.14.1833](https://doi.org/10.1001/jama.289.14.1833)] [Medline: [12684363](#)]
11. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. *Nicotine Tob Res* 2009 Sep;11(9):1025-1034. [doi: [10.1093/ntr/ntp090](https://doi.org/10.1093/ntr/ntp090)] [Medline: [19640833](#)]
12. Mohr DC, Vella L, Hart S, Heckman T, Simon G. The effect of telephone-administered psychotherapy on symptoms of depression and attrition: a meta-analysis. *Clin Psychol (New York)* 2008;15(3):243-253. [doi: [10.1111/j.1468-2850.2008.00134.x](https://doi.org/10.1111/j.1468-2850.2008.00134.x)] [Medline: [21369344](#)]

13. Carlbring P, Bohman S, Brunt S, Buhrman M, Westling BE, Ekselius L, et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. *Am J Psychiatry* 2006 Dec;163(12):2119-2125. [doi: [10.1176/appi.ajp.163.12.2119](https://doi.org/10.1176/appi.ajp.163.12.2119)] [Medline: [17151163](https://pubmed.ncbi.nlm.nih.gov/17151163/)]
14. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009;11(2):e13 [FREE Full text] [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
15. Reigeluth CM, Carr-Chellman A. Instructional-Design Theories and Models. Volume III, 3rd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 2009.
16. Keller JM. First principles of motivation to learn and e3-learning. *Distance Educ* 2008;29(2):175-185. [doi: [10.1080/01587910802154970](https://doi.org/10.1080/01587910802154970)]
17. Wigfield A, Eccles JS. Expectancy-Value Theory of Achievement Motivation. *Contemp Educ Psychol* 2000 Jan;25(1):68-81. [doi: [10.1006/ceps.1999.1015](https://doi.org/10.1006/ceps.1999.1015)] [Medline: [10620382](https://pubmed.ncbi.nlm.nih.gov/10620382/)]
18. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
19. Dick W, Carey L, Carey J. The Systematic Design of Instruction. New York, NY: Longman; 2001.
20. Richey RC, Klein JD. Developmental research methods: creating knowledge from instructional design and development practice. *J Comput High Educ* 2005;16(2):23-38. [doi: [10.1007/BF02961473](https://doi.org/10.1007/BF02961473)]
21. Richey R, Klein JD. Design and Development Research: Methods, Strategies, and Issues. 1st edition. Mahwah, NJ: L Erlbaum Associates; 2007.
22. Wang F. Design-based research and technology-enhanced learning environments. *Educ Technol Res Dev* 2005;53(4):5-23. [doi: [10.1007/BF02504682](https://doi.org/10.1007/BF02504682)]
23. Cobb P, Confrey J, diSessa A. Design experiments in educational research. *Educ Res* 2003;32(1):9-13. [doi: [10.3102/0013189X032001009](https://doi.org/10.3102/0013189X032001009)]
24. Gagne R, Wager W, Golas K, Keller J. Principles of Instructional Design. 5th edition. Belmont, CA: Thomson/Wadsworth; 2004.
25. Heinich R, Molenda M, Russell J, Smaldino S. The ASSURE model. In: Stollenwerk DA, editor. Instructional Media and Technologies for Learning. 7th edition. Upper Saddle River, NJ: Merrill/Prentice Hall; 2002.
26. Kirkpatrick D, Kirkpatrick J. Evaluating Training Programs: The Four Levels. 3rd edition. San Francisco, CA: Berrett-Koehler; 2006.
27. Smith PL, Ragan TJ. Instructional Design. 3rd edition. Hoboken, NJ: J Wiley & Sons; 2005.
28. Rossett A. First Things Fast: A Handbook for Performance Analysis (Essential Knowledge Resource). Hoboken, NJ: Pfeiffer; 2009.
29. Morrison G, Ross S, Kemp J. Designing Effective Instruction. 7th edition. Hoboken, NJ: Wiley; 2010.
30. Schank R, Neaman A. Motivation and failure in educational simulation design. In: Forbus KD, Feltovich PJ, editors. Smart Machines in Education: The Coming Revolution in Educational Technology. 1st edition. Menlo Park, CA: AAAI Press; 2001:37-69.
31. Wiggins G. A true test: toward more authentic and equitable assessment. *Phi Delta Kappan* 1989 May;70(9):703-713.
32. Gupta K. A Practical Guide to Needs Assessment. 2nd edition. San Francisco, CA: Pfeiffer/Wiley; 2007.
33. Ritterband LM, Cox DJ, Walker LS, Kovatchev B, McKnight L, Patel K, et al. An Internet intervention as adjunctive therapy for pediatric encopresis. *J Consult Clin Psychol* 2003 Oct;71(5):910-917. [doi: [10.1037/0022-006X.71.5.910](https://doi.org/10.1037/0022-006X.71.5.910)] [Medline: [14516239](https://pubmed.ncbi.nlm.nih.gov/14516239/)]
34. Conyne RK, Clack RJ. Environmental Assessment and Design: A New Tool for the Applied Behavioral Scientist. New York, NY: Praeger; 1981.
35. Bandura A. Human agency in social cognitive theory. *Am Psychol* 1989 Sep;44(9):1175-1184. [Medline: [2782727](https://pubmed.ncbi.nlm.nih.gov/2782727/)]
36. Prochaska JO, DiClemente CC. The transtheoretical approach. In: Norcross JC, Goldfried MR, editors. Handbook of Psychotherapy Integration. New York, NY: Basic Books; 1992:300-334.
37. Miller W, Rollnick S. Motivational Interviewing. 2nd edition: Preparing People for Change. New York, NY: The Guilford Press; 2002.
38. Schank RC, Berman TR. The pervasive role of stories in knowledge and action. In: Green MC, Strange JJ, editors. Narrative Impact: Social and Cognitive Foundations. Mahwah, NJ: L Erlbaum Associates; 2002:287-313.
39. Winn W. Research into practice: current trends in educational technology research: the study of learning environments. *Educ Psychol Rev* 2002;14(3):331-351. [doi: [10.1023/A:1016068530070](https://doi.org/10.1023/A:1016068530070)]
40. Anderson JR, Douglass S, Qin Y. How should a theory of learning and cognition inform instruction? In: Healy AF, editor. Healy AF, editor. Experimental Cognitive Psychology and its Applications. Washington, DC: American Psychological Association; 2005:47-58.
41. Lockhart RS, Craik FI. Levels of processing: a retrospective commentary on a framework for memory research. *Can J Psychol* 1990;44(1):87-112.
42. Mayer RE. Multimedia Learning. New York, NY: Cambridge University Press; 2009.
43. Gagne RM. Conditions of Learning. 2nd edition. New York, NY: Holt, Rinehart & Winston; 1970.

44. Weston C, McAlpine L, Bordonaro T. A model for understanding formative evaluation in instructional design. *Educ Technol Res Dev* 1995;43(3):29-48. [doi: [10.1007/BF02300454](https://doi.org/10.1007/BF02300454)]
45. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977 Mar;84(2):191-215. [Medline: [847061](https://pubmed.ncbi.nlm.nih.gov/847061/)]
46. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000 Jan;55(1):68-78. [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
47. Schunk DH, Zimmerman BJ. Self-regulation and learning. In: Reynolds WM, Miller GE, editors. *Handbook of Psychology: Educational Psychology*. Volume 7. New York, NY: Wiley; 2003:59-78.
48. Briggs LJ, Wager WW. *Handbook of Procedures for the Design of Instruction*. Englewood Cliffs, NJ: Educational Technology Publications; 1981.
49. Kinzie MB. Instructional design strategies for health behavior change. *Patient Educ Couns* 2005 Jan;56(1):3-15. [doi: [10.1016/j.pec.2004.02.005](https://doi.org/10.1016/j.pec.2004.02.005)] [Medline: [15590217](https://pubmed.ncbi.nlm.nih.gov/15590217/)]
50. Rosenstock IM. Historical origins of the health belief model. *Health Educ Monogr* 1974;2(4):328-335.
51. Dearing JW. An exploratory tool for predicting adoption decisions. *Sci Commun* 1994;16(1):43-57. [doi: [10.1177/0164025994016001003](https://doi.org/10.1177/0164025994016001003)]
52. Clark RC. Four architectures of instruction. *Perform Improv* 2000;39(10):31-38. [doi: [10.1002/pfi.4140391011](https://doi.org/10.1002/pfi.4140391011)]
53. Mayer RE. Unique contributions of eye-tracking research to the study of learning with graphics. *Learn Instruct* 2010;20(2):167-171. [doi: [10.1016/j.learninstruc.2009.02.012](https://doi.org/10.1016/j.learninstruc.2009.02.012)]
54. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckman J, editors. *Action Control: From Cognition to Behavior*. New York, NY: Springer-Verlag; 1985:11-39.
55. The Design-Based Research Collective. Design-based research: an emerging paradigm for educational inquiry. *Educ Res* 2003;32(1):5-8. [doi: [10.3102/0013189X032001005](https://doi.org/10.3102/0013189X032001005)]

Abbreviations

ID: instructional design

iSHIFUp: Internet Skin Health Intervention For Targeted Ulcer Prevention

SHUTi: Sleep Healthy Using The Internet

Edited by D Mohr, A Caele, H Christensen; submitted 15.07.11; peer-reviewed by D Kavanagh, N Kassam-Adams; comments to author 30.08.11; revised version received 14.11.11; accepted 16.11.11; published 28.06.12.

Please cite as:

Hilgart MM, Ritterband LM, Thorndike FP, Kinzie MB

Using Instructional Design Process to Improve Design and Development of Internet Interventions

J Med Internet Res 2012;14(3):e89

URL: <http://www.jmir.org/2012/3/e89/>

doi: [10.2196/jmir.1890](https://doi.org/10.2196/jmir.1890)

PMID: [22743534](https://pubmed.ncbi.nlm.nih.gov/22743534/)

©Michelle M Hilgart, Lee M Ritterband, Frances P Thorndike, Mable B Kinzie. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Internet-Based Interventions to Promote Mental Health Help-Seeking in Elite Athletes: An Exploratory Randomized Controlled Trial

Amelia Gulliver¹, BA, BScPsych (Hons); Kathleen M Griffiths¹, PhD; Helen Christensen¹, PhD; Andrew Mackinnon², PhD; Alison L Cascar¹, PhD; Alison Parsons¹, BA(Hons); Kylie Bennett¹, BSC, BA(Hons); Philip J Batterham¹, PhD; Rosanna Stanimirovic³, PhD

¹Centre for Mental Health Research, The Australian National University, Canberra, ACT, Australia

²Orygen Youth Health Research Centre, University of Melbourne, Melbourne, VIC, Australia

³Mental Edge Consulting, Melbourne, VIC, Australia

Corresponding Author:

Amelia Gulliver, BA, BScPsych (Hons)
Centre for Mental Health Research
The Australian National University
Building 63 Eggleston Road
Canberra, ACT, 0200
Australia
Phone: 61 2 6125 9472
Fax: 61 2 6125 0733
Email: amelia.gulliver@anu.edu.au

Abstract

Background: Mental disorders are more common in young adults than at any other life stage. Despite this, young people have low rates of seeking professional help for mental health problems. Young elite athletes have less positive attitudes toward seeking help than nonathletes and thus may be particularly unlikely to seek help. Interventions aimed at increasing help-seeking in young elite athletes are warranted.

Objective: To test the feasibility and efficacy of three Internet-based interventions designed to increase mental health help-seeking attitudes, intentions, and behavior in young elite athletes compared with a control condition.

Methods: We conducted a randomized controlled trial (RCT) of three brief fully automated Internet-based mental health help-seeking interventions with 59 young elite athletes recruited online in a closed trial in Australia. The interventions consisted of a mental health literacy and destigmatization condition, a feedback condition providing symptom levels, and a minimal content condition comprising a list of help-seeking resources, compared with a control condition (no intervention). We measured help-seeking attitudes, intentions and behavior using self-assessed surveys. Participation was open to elite athletes regardless of their mental health status or risk of mental illness.

Results: Of 120 athletes initially agreeing to participate, 59 (49%) submitted a preintervention or postintervention survey, or both, and were included in the present study. Adherence was satisfactory, with 48 (81%) participants visiting both weeks of assigned intervention material. None of the interventions yielded a significant increase in help-seeking attitudes, intentions, or behavior relative to control. However, at postintervention, there was a trend toward a greater increase in help-seeking behavior from formal sources for the mental health literacy/destigmatization condition compared with control ($P = .06$). This intervention was also associated with increased depression literacy ($P = .003$, $P = .005$) and anxiety literacy ($P = .002$, $P = .001$) relative to control at postintervention and 3-month follow-up, respectively, and a reduction in depression stigma relative to control at postintervention ($P = .01$, $P = .12$) and anxiety stigma at 3-month follow-up ($P = .18$, $P = .02$). The feedback and help-seeking list interventions did not improve depression or anxiety literacy or decrease stigmatizing attitudes to these conditions. However, the study findings should be treated with caution. Due to recruitment challenges, the achieved sample size fell significantly short of the target size and the study was underpowered. Accordingly, the results should be considered as providing preliminary pilot data only.

Conclusions: This is the first RCT of an Internet-based mental health help-seeking intervention for young elite athletes. The results suggest that brief mental health literacy and destigmatization improves knowledge and may decrease stigma but does not increase help-seeking. However, since the trial was underpowered, a larger trial is warranted.

Trial Registration: 2009/373 (www.clinicaltrials.gov ID: NCT00940732), cited at <http://www.webcitation.org/5ymsRLy9r>.

(*J Med Internet Res* 2012;14(3):e69) doi:[10.2196/jmir.1864](https://doi.org/10.2196/jmir.1864)

KEYWORDS

Mental health; help-seeking; elite athletes

Introduction

Mental disorders contribute significantly to worldwide disease burden [1,2]. Depression and anxiety in particular are common, with an estimated prevalence of 6%–18% in high-income countries globally [3–5]. Mental disorders are more common in young adults than at any other stage of the lifespan [3]. Despite this, young people have low rates of seeking professional help for mental health problems [6]. Research indicates that athletes may experience a prevalence of symptoms of depression similar to that of the general population [7] and have less positive attitudes toward seeking help than nonathletes [8]. Since attitudes are thought to influence help-seeking in the general population [6,9,10] and this age group is already at risk, young elite athletes may be particularly vulnerable to not seeking help.

Research has examined the reasons why young people do not seek professional help for mental disorders. A systematic review investigating young people's perceived barriers to help-seeking found that stigma and embarrassment, problems recognizing symptoms (poor mental health literacy), and a preference for self-reliance were key barriers to seeking help [11]. Factors that might facilitate help-seeking have been comparatively underresearched. However, there has been some evidence that positive past experiences, and social support and encouragement from others may facilitate help-seeking [11]. These factors are said to be effective because they increase mental health literacy and reduce stigma.

Several studies have evaluated interventions designed to improve help-seeking. Previous randomized controlled trials (RCTs) have used information-based methods of targeting help-seeking via video [12], written and verbal information [13–15], and the Internet [16,17]. These interventions have demonstrated an increase in positive attitudes to help-seeking [12–14], willingness to seek help [15], help-seeking beliefs [17], and, less frequently, behavior [16]. Only one trial has been conducted within the sports field. It found that providing information about the benefits of sports psychology in an interview improved athletes' attitudes to consulting a sports psychologist [13]. There have been no other studies of the efficacy of interventions specifically designed to encourage help-seeking behavior in athletes.

It is not clear from these previous studies what might be the best approach to encouraging help-seeking among youth or particularly elite athletes. Rickwood et al [18] conceptualized seeking professional help as a 4-step process beginning with the individual becoming aware of a problem, then expressing the problem to others, followed by identifying appropriate and accessible sources of help, and then seeking out a health

professional and disclosing his or her problems. However, it is not clear which of or how many of these stages are critical to encourage professional help-seeking. Consequently we designed three different interventions, each of which focused on one or more of these stages. The first, a mental health literacy/destigmatization intervention, addressed three of the four stages. First, information was provided to increase the participant's awareness of a potential problem. Second, the intervention incorporated messages designed to decrease stigma in order to facilitate the expression of a problem. Third, a list of help-seeking sources was provided to assist with the identification of appropriate resources. The second intervention provided feedback about symptoms. The feedback intervention thus focused on the first stage of the Rickwood model to increase awareness of a potential problem. Participants were explicitly provided with information about their levels of symptoms relative to other individuals of a similar age. This intervention also provided information about help-seeking sources. The third intervention was a help-seeking list intervention. This intervention consisted of a list of help-seeking resources to determine whether providing this information alone was sufficient to encourage the participants to seek help. The control condition, where participants received no intervention, was introduced to control for the effect of participating in a research trial and the elapse of time [17].

We hypothesized that, compared with the control group, participants in all three intervention conditions would demonstrate at postintervention and 3-month follow-up improved professional help-seeking attitudes and intentions, and help-seeking from professional sources. Given differences in content and level of intensity of the interventions, we predicted the effects would be strongest in the mental health literacy/destigmatization condition, followed by the feedback condition, and then the help-seeking list condition. Because the help-seeking list was used in all three active interventions, it was possible to compare the relative efficacy of the mental health literacy and feedback components of the first two interventions over and above the effects of the list alone. In addition to differences in help-seeking, we hypothesized that those in the mental health literacy/destigmatization condition would demonstrate increased mental health literacy and decreased stigma compared with those in the control condition.

The three active interventions were delivered through the Internet. Internet-delivered interventions are available 24 hours a day, 7 days a week, can be accessed anonymously, are cost effective [19,20], and can be widely distributed [21,22]. Despite these benefits, only two trials have evaluated the utility of an Internet-based format for encouraging help-seeking. This is the

first RCT to examine the feasibility and efficacy of Internet-based help-seeking interventions in elite athletes.

Methods

The Elite Athlete Mental Health Strategy (TEAMS) project was undertaken in two stages. Stage 1 involved an online survey of elite athletes. This survey incorporated self-report measures of a range of mental health symptoms, and demographic and other participant attributes. Following this, we invited the athletes to participate in stage 2 of the RCT. Inclusion criteria for both stages of the study were being aged 18 years or older and being an elite athlete as defined by their level of competition (Olympic or Paralympic, professional, or state-, national-, or international-level athletes). An implicit inclusion criterion was that the participants should be Internet and computer literate. Ethics approval for the study was granted by both the Australian Institute of Sport (AIS) ethics committee and The Australian National University Human Research Ethics Committee (ANU HREC 2009/373).

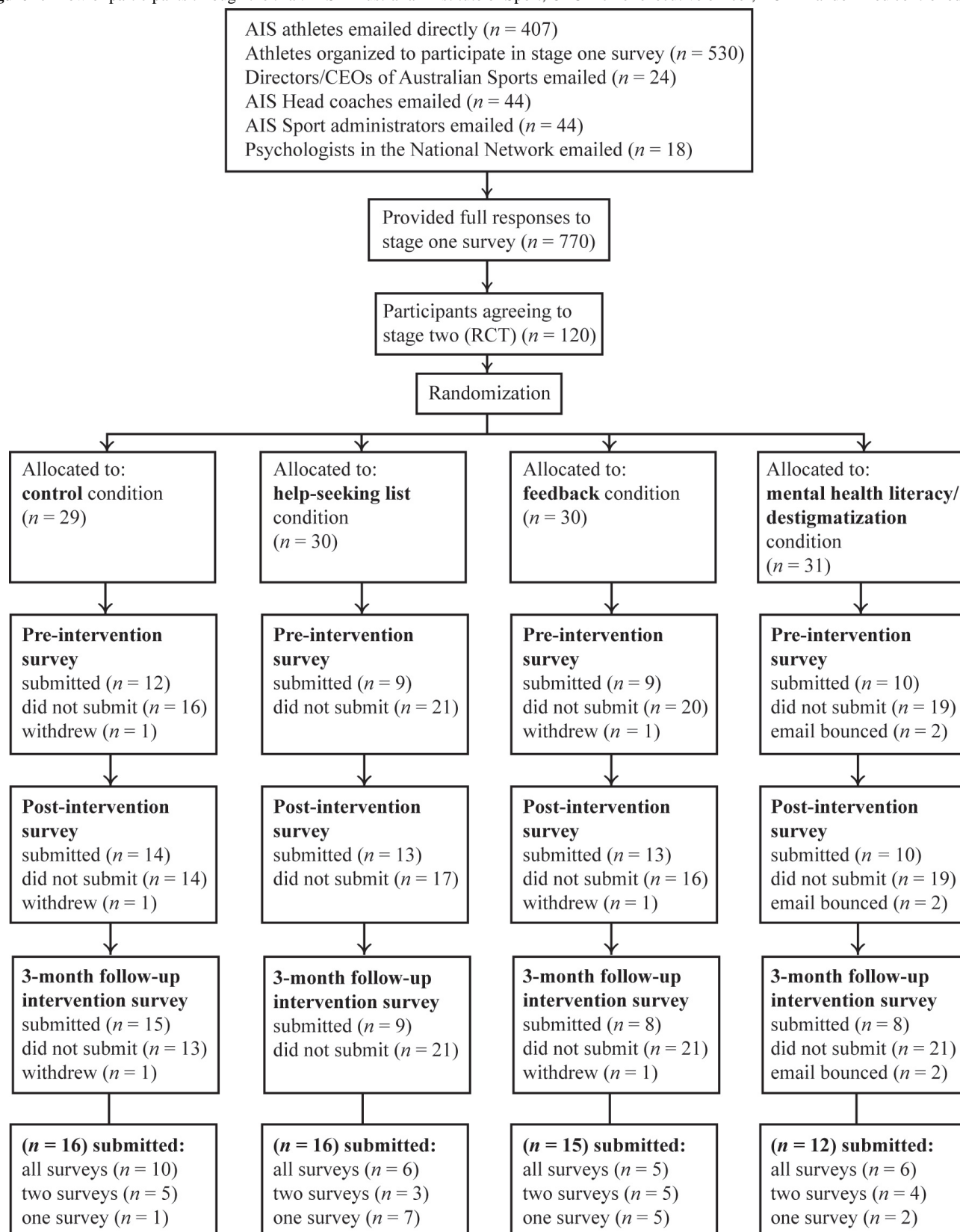
Participants and Recruitment

Participants were elite athletes from the AIS and from other national sporting organizations in Australia. They were recruited primarily via emails distributed by the Director of the AIS and through direct recruitment with elite sporting clubs. An initial email invitation containing an embedded link to the information and consent page of the project website was sent by the Director of the AIS in November 2009 to directors and chief executive officers of elite sports organizations around Australia, as well as to all AIS head coaches and administrators, and psychologists from the Australian Psychological Society College of Sport Psychologists. The distribution list was compiled by the AIS. However, this recruitment strategy was of limited effectiveness. Accordingly, we sent a follow-up email to directors and coaches

requesting that they distribute the aforementioned emails directly to athletes. However, the number of participants remained low. Therefore, we implemented a third recruitment strategy. This involved sending an initial and a follow-up advertisement-style email from the AIS Director with AIS branding, in both HTML and plain text, directly to all AIS athletes aged 18 years or older ($n = 407$) during March 2010 (recruitment wave 1). In addition, one organization arranged to send a text message to their athletes' mobile phones (January 2010) to indicate that they had been sent an email. The final recruitment strategy was to organize a large number of athletes to participate in a group setting at a sports club site using computer facilities. This resulted in a substantial increase in participants ($n = 530$; see [Figure 1](#)) into stage 1 of the project from November 2010 to February 2011 (recruitment wave 2), but a limited number of athletes agreeing to participate in the stage 2 trial.

We offered a random prize draw incentive of an iPhone 3GS (wave 1) and an iPhone 4 (wave 2) to participants who completed the postintervention survey. Incentives have previously been demonstrated to increase participants' retention rates, training acceptance, and self-assessed effectiveness of Internet-based interventions [23].

[Figure 1](#) presents the flow of participants through the trial. At the conclusion of the stage 1 survey, participants were provided with information about the RCT, and those who agreed to participate provided their email address and were automatically randomly assigned to an intervention arm. Overall, 120 athletes were randomly assigned to the following conditions: control, help-seeking list, feedback, and mental health literacy/destigmatization. A total of 59 participants (16 men, 43 women) completed at least a preintervention or a postintervention survey and were included in the present study results. Participants' age ranged from 18 to 48 years, and the mean age was 25.5 years (median 24.5).

Figure 1. Flow of participants through the trial. AIS = Australian Institute of Sport; CEO = chief executive officer; RCT = randomized controlled trial.

Treatment Allocation

In an email message, participants were provided with a link to an anonymous online survey comprising demographic and other mental health measures (stage 1). This survey was located on

the AIS-branded Web portal created for the project. One sporting organization arranged for their athletes to complete this initial survey at a set location using laptops (n = 530; see Figure 1). Immediately after completing the stage 1 survey, participants who indicated that they were interested in participating in stage

2 of the project (RCT) were redirected to a webpage on the project website where they could register their email address to be involved in the trial. Participants were then randomly assigned to the different conditions using predefined automated computerized block randomization with a block size of 16. The automated computerized system was set up by researchers not involved in the day-to-day management of the study.

Data Collection

The intervention and measurement points were spaced 1 week apart and comprised the stage 1 survey (demographics, week 1), preintervention survey (week 2), intervention week 1 (week 3), intervention week 2 (week 4), and postintervention survey (week 5). An automated email was sent to the participants each week with an embedded link to the self-assessed surveys or intervention content. A single reminder email was sent if, after 4 days, participants had not yet accessed their assigned survey or intervention webpage by clicking on the embedded link. Two further follow-up survey emails were sent automatically at 3 and 6 months after the preintervention survey. Each survey took approximately 15–20 minutes to complete. Finally, a debriefing email providing information about the purpose of the study was sent 1 month after the 6-month follow-up email.

Online Surveys

The information and consent webpage apprised participants of what each stage of the study would involve (including survey time), as well as all requirements listed by the ethics review boards, including when and where data would be securely stored, the investigators and their contact details, and the general purpose of the study. We tested the online surveys by the researchers before deployment and provided them on an open access website; however, we emailed links to this website only to sports managers and elite athletes. Tokens enabled the unique identification of each participant, and each token was embedded into the link to their unique survey. Each page of the survey comprised items from a single measure. Each of the measurement points comprised 10 pages (excluding the stage 1 survey, which comprised 23 pages). All items were mandatory, although participants could review and change their answers before submitting them. In addition, the help-seeking behavior items of the General Help-Seeking Questionnaire (GHSQ) used adaptive questioning, whereby the qualitative items would appear only if the help-seeking sources were selected. Data were automatically captured by the online survey platform, and we used data from all surveys regardless of completeness; thus, participant numbers may differ for each of the outcome variables.

Interventions

Although participants could not be blinded to their assigned condition, they were informed on the information and consent page that they would be asked to undertake “some online activities” regarding the “usefulness of different types of health information.” This general description reduced the likelihood that participants would be aware of the anticipated effectiveness of their assigned condition.

Intervention content was spaced approximately evenly over 2 weeks and consisted of a series of linear webpages that varied

in length depending on the condition type. At the end of each intervention week, the participants were provided with a list of sources of help and resources for mental health problems, which were identical across all intervention conditions at both week 1 and week 2. Sources included face-to-face, telephone, and emergency sources (week 1), and an additional hyperlinked list of online sources for cognitive behavioral therapy and bibliotherapy (week 2). All intervention conditions included an introduction to the week’s activities and a final page congratulating the participant for completing the activities. We made no changes to the intervention content during the trial. The specific content of each of the intervention conditions is described below.

Mental Health Literacy/Destigmatization

This condition comprised written material created by one of the researchers (KG) delivered on 34 brief linear webpages. It was designed to increase mental health literacy and decrease stigma, specifically targeting depression and anxiety. The first week contained 19 pages of information about the prevalence and risk of mental disorders; disability and symptoms of depression; a matrix of evidence-based treatments demonstrating the most effective treatments for depression; detailed information about the two most effective depression treatments—antidepressants and cognitive behavioral therapy; a consideration of four myths about depression aimed specifically at reducing stigma; and help-seeking sources (as listed above). The second week was provided in a similar format, this time focusing on generalized anxiety disorder, with 15 pages of brief information about social anxiety, panic, posttraumatic stress, and obsessive compulsive disorders, as well as specific phobia.

Feedback

The second condition comprised two interactive quizzes providing tailored feedback to the participant about his or her level of depression (week 1) and anxiety (week 2) using the Goldberg Depression and Anxiety scales [22]. Participants received one of four possible feedback pages corresponding to their symptom score. We derived norms for this feedback from the MoodGYM clinician’s manual [24], which was based on data from a sample of 2404 Australian community-dwelling young adults aged 20–24 years (see [25] for details of sample). It has been suggested that providing feedback is the most efficacious of different types of tailoring mechanisms in implementing behavioral change in smoking cessation [26], and it also appears promising in influencing other health behaviors [27–29]. Thus, we designed this condition to provide an objective measure of depression and anxiety symptoms to allow the participants to assess their need to seek help. The content of this condition was based on the FRAMES framework for effective brief interventions for behavioral change [30]. FRAMES is an acronym for the inclusion of feedback (a form of assessment), responsibility (advising that the behavior is the individual’s responsibility and choice), advice (providing verbal or written advice to change), menu (providing an array of alternative strategies for change), empathy (adopting a warm, reflective, and understanding approach), and self-efficacy (emphasizing the individual’s self-efficacy for change). The

condition comprised six pages each week for week 1 (depression) and week 2 (anxiety).

Help-Seeking List

This condition comprised three pages each week, including an introduction, help-seeking source page (see above for details), and a congratulatory page.

Control

Participants in the control condition received emails to the online measurement surveys only.

Measures

We collected a large number of demographic and other characteristics and symptoms of mental disorders in the stage 1 survey. However, we present only key, relevant demographic data in the current paper. These included gender, age, highest level of education, level of competition, sports type, psychological distress as measured by the 10-item Kessler Psychological Distress Scale (K10) [31], and prior help-seeking as measured by the prior counseling measure of the GHSQ [32].

Primary Outcome Measures

The primary outcome measures were help-seeking attitudes, intentions, and behaviors. Attitudes were measured by the Attitudes Toward Seeking Professional Psychological Help-Short Form (ATSPPH-SF) [33]. The ATSPPH-SF is a 10-item scale measuring attitudes toward seeking psychological help from a professional. Items on the scale include “I would want to get psychological help if I were worried or upset for a long period of time.” Participants indicated their agreement with each statement on a 4-point Likert-type scale ranging from 0 (strongly disagree), 1 (partly disagree), 2 (partly agree), to 3 (strongly agree). This response set was different from the original (strongly disagree and strongly agree should be disagree and agree, respectively) due to a transcription error. Responses were coded from 0 to 3 and summed to calculate a total score, after reverse scoring negatively worded items. Total scores ranged from 0 to 30, with higher scores indicating more positive attitudes to professional psychological help-seeking. Research has previously reported good internal consistency for the ATSPPH-SF in studies of university students (alpha range of .72–.85) [33–36] and medical patients (alpha = .78) [34], as well as adequate validity [33,34] and acceptable test–retest reliability ($r = .80$) [33]. In the present study the internal reliability using Cronbach alpha was .69 ($n = 40$) at preintervention. Test–retest reliability was moderate between preintervention and 4 weeks postintervention for the control condition ($r = .64$; $n = 10$, $P = .045$).

We measured help-seeking intentions by the intentions scale of the GHSQ [37]. The scale used in the present study consisted of 11 items measuring the respondent’s intentions to seek help for a “personal or emotional” problem. The respondent was first asked “If you were having a personal or emotional problem, how likely is it that you would seek help from each of the following sources?” The scale comprised potential sources of formal and informal help (eg, “friend (not related to you)” or “Doctor/GP”) and included an option that they would not seek help from anyone. Each source of help was rated on a 7-point

Likert-type scale ranging from 1 (extremely unlikely) to 7 (extremely likely). Item scores ranged from 1 to 7, with higher scores indicating higher intentions to seek help from that particular source. The GHSQ contains two subscales [38]. The first comprises the mean of formal sources (“Doctor/GP” and “Mental Health Professional (eg, school counselor, psychologist, psychiatrist)”, and the second comprises the mean of informal sources (“Intimate partner(eg, significant boyfriend or girlfriend, husband, wife)”, “Friend (not related to you)”, “Mother”, “Father”, and “Other relative/family member”).

The intentions scale of the GHSQ has adequate psychometric properties. Research has previously reported broadly adequate internal consistency in studies of high school students aged 12–19 years for personal or emotional problems (alpha range of .70–.90) [32,39] and university students (alpha = .67) [37], as well as adequate validity [32] and very good test–retest reliability ($r = .86$) [32]. Internal consistency at preintervention in the present study was adequate for formal sources (alpha = .77, $n = 37$), although lower than ideal for informal sources (alpha = .57, $n = 37$). Test–retest reliability at preintervention and 4 weeks postintervention for the control condition was low for formal sources ($r = .42$; $n = 9$, $P = .26$) and high for informal sources ($r = .91$; $n = 9$, $P = .001$).

Finally, we measured help-seeking behavior by the proportion of participants who reported seeking help from a formal source in the past 3 weeks (since the beginning of the intervention), using the Actual Help-Seeking Questionnaire (AHSQ) [32,40]. This questionnaire consists of 11 items measuring the respondents’ reported help-seeking behavior in the past 2 weeks for a “personal or emotional” problem. Participants were provided a list of help-seeking sources (matched to the list for the intentions scale of the GHSQ reported above [33]) and asked to indicate which, if any, they had “gone to for advice or help in the past 3 weeks.” Participants were also asked to briefly indicate the type of problem they “went to them about.” Participants were also able to indicate that they had not spoken with anyone about their problems or that they had not experienced any problems. As per above, the AHSQ contains two subscales, one based on the formal help-seeking items and the other on the informal items.

Secondary Outcome Measures

The secondary outcomes were mental health literacy and stigma.

We assessed mental health literacy specific to depression and anxiety on the Depression Literacy questionnaire (D-Lit) [41] and a new measure, the Anxiety Literacy questionnaire (A-Lit, developed by KG). These measures each consisted of 22 items, with items consisting of a statement assessing the respondents’ knowledge about depression and anxiety, respectively. For each statement, respondents selected what they believed to be the correct response from three possible choices (true, false, or I don’t know). There is a mix of true and false items in each scale. For example, “Loss of confidence and poor self-esteem may be a symptom of depression” (D-Lit; true), and “People with anxiety disorder often hear voices that are not there” (A-Lit; false). Respondents scored 1 point for each correct answer, and total scores ranged from 0 to 22 with higher scores indicating higher literacy for depression (D-Lit) or anxiety (A-Lit).

Psychometric properties are available for the D-Lit. A previous study reported very good internal consistency for the Greek and Italian versions of the D-Lit ($\alpha = .88$ and $\alpha = .92$, respectively) [42], as well as adequate validity and very good test–retest reliability ($r = .91$) [42]. Internal consistency for the D-Lit in the present study was adequate with a Cronbach alpha coefficient of $.70$ ($n = 40$). Test–retest reliability for preintervention and 4 weeks postintervention for the control condition was also adequate ($r = .71$; $n = 12$, $P = .02$). For the A-Lit, internal consistency was acceptable with a Cronbach alpha coefficient of $.76$ ($n = 40$), and test–retest reliability for the control condition was very good ($r = .83$; $n = 12$, $P = .003$).

Similarly, we used two scales to measure personal stigma toward depression or anxiety: the Personal Stigma items of both the Depression Stigma Scale (DSS) [41] and the Generalised Anxiety Stigma Scale (GASS) [43]. The DSS comprises 9 items and the GASS comprises 10 items measuring the respondents' own stigmatizing attitudes toward depression and anxiety, respectively. Items consist of statements such as "Depression is not a real medical illness" (DSS) and "An anxiety disorder is a sign of personal weakness" (GASS). Respondents indicate their agreement with each statement on a 5-point Likert-type scale ranging from 0 (strongly disagree) to 4 (strongly agree). Responses are summed to calculate a total score, which ranges from 0 to 36 for the DSS, and from 0 to 40 for the GASS. Higher scores on each scale indicate more stigmatizing attitudes toward depression (DSS) or anxiety (GASS). Previous research has reported good internal consistency for the DSS in community studies of adults (α range of $.76$ – $.82$) [41,44], as well as acceptable validity and test–retest reliability ($r = .66$) [41]. In the present study the internal reliability for the DSS was good at preintervention with a Cronbach alpha coefficient of $.77$ ($n = 40$). Test–retest reliability for the DSS was high between preintervention and 4 weeks postintervention for the control condition ($r = .86$; $n = 12$, $P = .001$). The GASS scale has also been shown to have good internal consistency in a community study of adults ($\alpha = .86$) [43], as well as acceptable validity and moderate test–retest reliability ($r = .58$) [43]. In the present study the internal consistency for the GASS was high at preintervention with a Cronbach alpha coefficient of $.90$ ($n = 39$). In addition, this sample demonstrated high test–retest reliability for the GASS between preintervention and 4 weeks postintervention for the control condition ($r = .91$; $n = 12$, $P < .001$).

Statistical Methods

We assessed the study outcomes within an intention-to-treat framework involving the analysis of data across the first three measurement occasions. The effect of the interventions was examined using mixed-models repeated-measures analysis of variance for longitudinal continuous data using SPSS 19.0 for

Windows (IBM Corporation, Somers, NY, USA). We used mixed-model logistic regression to analyze longitudinal binary data, using the xtlogit procedure of Stata version 10 for Windows (StataCorp LP, College Station, TX, USA). The within-groups factor was measurement occasion, and the between-groups factor was condition type. Mixed models were used to include all available data, including that from participants who dropped out or did not provide data at pre- or postintervention, under the assumption that data were missing at random [45,46]. The analyses focused on the interactions between the main effects of condition type and measurement occasion. We used planned contrasts to assess change over time between each intervention condition and all other conditions including control. Postintervention contrasts assessed change over time from preintervention to postintervention, and 3-month follow-up contrasts assessed change from preintervention to 3-month follow-up. To evaluate the magnitude of the effects given the small sample size, we estimated between-group effect sizes between each intervention and the control condition with observed means at postintervention and 3-month follow-up using Hedges' g [47] to account for the small sample size. For consistency in interpretation, in scales where a decreasing score indicated an advantage for the intervention group, we reversed the formula so that all effect sizes were positive.

Power Calculations

The target sample size was 500 participants (125 per condition), which we chose to detect an effect size of 0.25 in help-seeking attitudes between a comparison group and one of the active conditions with 80% power at $\alpha = .05$, assuming a correlation between pre- and postintervention of $.5$. This sample size was appropriate given the low to moderate effect sizes for help-seeking measures previously detected in RCTs of help-seeking (Cohen d range of -0.08 to 0.56) [12–17]. The achieved sample size ($n = 59$) would have required an effect size of approximately 0.75 standard deviations for 80% power at $\alpha = .05$, and with a correlation between pre- and postintervention of $.5$. The present study sample was therefore underpowered and the results should be considered as preliminary pilot data.

Results

Table 1 presents sample characteristics. The age of participants ranged from 18 to 48 years. Participants were from the sports of cricket ($n = 14$, 24%), soccer ($n = 9$, 15%), rowing ($n = 5$, 8%), sailing ($n = 4$, 7%), field hockey ($n = 4$, 7%), and multiple (14) other sports ($n = 23$, 39%). The majority of participants were Olympic athletes ($n = 10$, 17%) or international-level athletes ($n = 28$, 48%), while the remainder ($n = 21$, 36%) were competing at national state, professional, age group international, Paralympic, or age group national state levels.

Table 1. Characteristics of participants included in the study.

Characteristic	Intervention arm				Total (n = 59)
	Mental health literacy/destigmatization (n = 12)	Feedback (n = 15)	Help-seeking list (n = 16)	Control (n = 16)	
Age (years), mean (SD)	24.83 (5.32)	25.80 (4.51)	25.44 (7.10)	25.50 (5.68)	25.42 (5.64)
Gender, n (%)					
Male	2 (17%)	5 (33%)	5 (31%)	4 (25%)	16 (27%)
Female	10 (83%)	10 (67%)	11 (69%)	12 (75%)	43 (73%)
Highest level of postsecondary/tertiary education, n (%)					
Trade/apprenticeship	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Other certificate	0 (0%)	3 (20%)	5 (31%)	2 (13%)	10 (17%)
Associate or undergraduate diploma	0 (0%)	2 (13%)	1 (6%)	0 (0%)	3 (5%)
Bachelor's degree	5 (42%)	5 (33%)	5 (31%)	6 (38%)	20 (36%)
Higher degree	1 (8%)	1 (7%)	0 (0%)	3 (19%)	5 (9%)
Other	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
None	6 (50%)	4 (27%)	5 (31%)	5 (31%)	20 (34%)
K10 ^a score, mean (SD)	16.17 (6.55)	16.47 (5.82)	14.00 (1.79)	15.19 (4.76)	15.39 (5.64)
Previously sought help, n (%)	8 (67%)	10 (67%)	9 (56%)	7 (44%)	34 (58%)
Rated helpfulness of the visits, mean (SD)	3.75 (0.89)	3.44 (1.51) ^b	3.11 (1.45)	4.17 (0.98) ^b	3.56 (1.27)

^a 10-item Kessler Psychological Distress Scale.

^b Missing data on this item for one respondent from the feedback condition and one from the control condition.

Table 2 presents the preintervention comparisons between conditions. There were no significant differences between the intervention conditions for any of these variables. In addition, at preintervention there were no significant differences between

trial conditions in the proportion of participants who were of Olympic and international (including Paralympic and age group international) level ($\chi^2_{3=0.3}$, $P = .98$).

Table 2. Preintervention comparisons between conditions.

Characteristic	Test of significance
Age in years	$F_{3,55} = 0.06, P = .98$
Gender (male/female)	$\chi^2_3 = 1.1, P = .79$
Highest level of postsecondary/tertiary education (those with a higher-level education/those without)	$\chi^2_3 = 1.8, P = .65$
General psychological distress (K10 ^a)	$F_{3,55} = 0.77, P = .51$
Prior counseling experience (yes/no)	$\chi^2_3 = 2.2, P = .55$
Rated helpfulness of the visits	$F_{3,28} = 0.91, P = .45$
Help-seeking attitudes (ATSPPH-SF ^b)	$F_{3,36} = 0.59, P = .63$
Help-seeking intentions (GHSQ^c)	
Informal sources	$F_{3,33} = 1.24, P = .31$
Formal sources	$F_{3,33} = 0.70, P = .56$
Help-seeking behavior (AHSQ^d)	
Informal sources	$\chi^2_3 (n = 40) = 3.7, P = .33$
Formal sources	$\chi^2_3 (n = 40) = 1.0, P = .85$
Depression literacy (D-Lit ^e)	$F_{3,36} = 0.75, P = .53$
Anxiety literacy (A-Lit ^f)	$F_{3,36} = 0.48, P = .69$
Depression stigma (DSS ^g)	$F_{3,36} = 0.80, P = .50$
Anxiety stigma (GASS ^h)	$F_{3,35} = 1.34, P = .28$

^a 10-item Kessler Psychological Distress Scale.

^b Attitudes Toward Seeking Professional Psychological Help-Short Form.

^c General Help-Seeking Questionnaire.

^d Actual Help-Seeking Questionnaire.

^e Depression Literacy questionnaire.

^f Anxiety Literacy questionnaire.

^g Depression Stigma Scale.

^h Generalised Anxiety Stigma Scale.

There was no significant difference between conditions in those missing at both pre- and postintervention from among those who agreed to participate in the trial ($N = 120, \chi^2_{3=1.8}, P = .63$). Nor was there a significant difference between conditions in those missing at postintervention ($n = 40, \chi^2_{3=0.9}, P = .92$) or at 3-month follow-up ($n = 40, \chi^2_{3=4.6}, P = .22$) from among those present at preintervention.

We conducted logistic regression analyses to assess whether there were any significant predictors of agreeing to participate in the trial ($n = 120$) from those completing the stage 1 survey ($N = 770$); the completion of any survey ($n = 59$) from those agreeing to participate ($N = 120$); and the completion of the postintervention survey ($N = 50$) from those who completed the preintervention survey ($N = 40$). Agreeing to participate in the trial was significantly related to gender, with men less likely than women to agree (odds ratio [OR] 0.14, 95% confidence interval [CI] 0.08–0.22, $P < .001$); age, with older participants more likely than younger participants to agree (OR 1.08, 95%

CI 1.03–1.13, $P = .001$); prior counseling experience, with those who had *not* received previous counseling more likely to agree (OR 1.74, 95% CI 1.09–2.76, $P = .02$); and general psychological distress, with those with higher K10 scores more likely to agree (OR 1.05, 95% CI 1.01–1.09, $P = .02$). Agreement to participate in the trial was not related to highest level of education. Missingness for both surveys (not submitting either the pre- or postintervention survey after agreeing to participate) was significantly related to gender only, with men more likely than women to be missing (OR 12.90, 95% CI 4.79–34.74, $P < .001$). Missingness for both surveys was not related to condition, age, highest level of education, prior counseling experience, or general psychological distress. Missingness at postintervention (not submitting a postintervention survey) or at 3-month follow-up (not submitting a 3-month follow-up survey) from those present at preintervention was not related to condition, age, gender, posttertiary education, prior counseling experience, or general psychological distress.

Intervention Adherence

We used unique identifiers (tokens) to track each participant's use of the intervention materials. Of the 43 study participants from the intervention conditions, 41 (95%) visited at least 1 week's online program, 35 (81%) visited both weeks' online programs, and 2 (5%) visited neither. The intervention condition

had no effect on the number of online programs visited ($F_{2,39} = 0.13$, $P = .88$).

Primary Intervention Effects

Table 3 presents the observed means, standard deviations, and proportions for the primary outcome variables.

Table 3. Observed means, standard deviations, and proportions for the primary outcome measures at each measurement occasion for the trial conditions.

Measure	Measurement occasion					
	Preintervention		Postintervention		3-month follow-up	
	n	Data	n	Data	n	Data
Help-seeking attitudes (ATSPPH-SF^a), mean (SD)						
Mental health literacy/destigmatization	10	20.30 (3.34)	10	22.40 (3.34) ^b	8	20.63 (4.27)
Feedback	8	20.67 (3.00)	12	20.67 (4.19)	8	22.63 (4.00)
Help-seeking list	8	20.22 (2.73)	12	20.92 (3.23)	9	22.44 (3.47)
Control	11	21.92 (4.10)	14	21.14 (5.26)	15	21.27 (5.16)
Help-seeking intentions (GHSQ^c) – formal sources, mean (SD)						
Mental health literacy/destigmatization	10	4.05 (1.52)	10	4.10 (1.29)	8	4.19 (1.60)
Feedback	8	3.63 (1.85)	12	3.92 (1.22)	8	4.38 (1.66)
Help-seeking list	8	4.63 (1.22)	12	4.04 (0.92)	9	4.33 (1.32)
Control	11	4.50 (1.69)	14	3.79 (1.81)	15	4.10 (1.84)
Help-seeking intentions (GHSQ) – informal sources, mean (SD)						
Mental health literacy/destigmatization	10	5.00 (0.92)	10	5.12 (0.81)	8	4.78 (0.93)
Feedback	8	4.58 (2.05)	12	4.27 (1.49)	8	4.43 (0.88)
Help-seeking list	8	5.55 (1.18)	12	5.05 (1.29)	9	5.38 (1.09)
Control	11	5.58 (0.86)	14	5.17 (1.21)	15	5.15 (1.12)
Help-seeking behavior (GHSQ) – sought formal help, n (%)						
Mental health literacy/destigmatization	10	2 (20%)	10	5 (50%)	8	3 (38%)
Feedback	9	1 (11%)	12	1 (8%)	8	3 (38%)
Help-seeking list	9	1 (11%)	12	2 (17%)	9	4 (44%)
Control	12	3 (25%)	14	1 (7%)	15	3 (20%)
Help-seeking behavior (GHSQ) – sought informal help, n (%)						
Mental health literacy/destigmatization	10	6 (60%)	10	7 (70%)	8	6 (75%)
Feedback	9	8 (89%)	12	7 (58%)	8	6 (75%)
Help-seeking list	9	5 (56%)	12	8 (67%)	9	7 (78%)
Control	12	6 (50%)	14	9 (64%)	15	12 (80%)

^a Attitudes Toward Seeking Professional Psychological Help-Short Form.

^b Significant pattern of change from pre- to postintervention vs feedback, $P = .04$.

^c General Help-Seeking Questionnaire.

Help-Seeking Attitudes

The overall interaction between condition and measurement occasion for help-seeking attitudes was not significant ($F_{6,68.92}$

$= 1.64$, $P = .15$). None of the conditions, including the control, had significant changes from pre- to postintervention or from preintervention to 3-month follow-up. However, there was an improvement from pre- to postintervention in attitudes in the

mental health literacy/destigmatization condition ($P = .04$) that was significant compared with the feedback condition only, for which the estimated marginal means for attitudes decreased at postintervention. Hedges' g between-group effect sizes at postintervention for the intervention conditions compared with the control condition were 0.28 (95% CI -0.54 to 1.09) for the mental health literacy/destigmatization condition, -0.12 (95% CI -0.91 to 0.67) for the feedback condition, and -0.05 (95% CI -0.82 to 0.72) for the help-seeking list condition. At 3-month follow-up, effect sizes compared with the control for help-seeking attitudes were as follows: mental health literacy/destigmatization condition ($g = -0.13$, 95% CI -0.99 to 0.73), feedback condition ($g = 0.29$, 95% CI -0.58 to 1.15), and help-seeking list condition ($g = 0.26$, 95% CI -0.57 to 1.09).

Help-Seeking Intentions

The overall interaction between condition and measurement occasion was not significant for either formal ($F_{3,35.45} = 0.55$, $P = .65$) or informal help-seeking intentions ($F_{3,26.29} = 2.21$, $P = .11$). Between-group effect sizes at postintervention for formal and informal intentions between the control condition and the interventions conditions were as follows: mental health literacy/destigmatization condition (formal, $g = 0.19$, 95% CI -0.62 to 1.01 ; informal, $g = -0.05$, 95% CI -0.86 to 0.76), feedback condition (formal, $g = 0.08$, 95% CI -0.69 to 0.85 ; informal, $g = -0.67$, 95% CI -1.47 to 0.12), and help-seeking list condition (formal, $g = 0.17$, 95% CI -0.60 to 0.94 ; informal, $g = -0.10$, 95% CI -0.87 to 0.67). At 3-month follow-up, effect sizes compared with the control for help-seeking intentions were as follows: mental health literacy/destigmatization condition (formal, $g = 0.05$, 95% CI -0.81 to 0.91 ; informal, $g = -0.35$, 95% CI -1.22 to 0.51), feedback condition (formal, g

$= 0.16$, 95% CI -0.70 to 1.02 ; informal, $g = -0.69$, 95% CI -1.58 to 0.19), and help-seeking list condition (formal, $g = 0.14$, 95% CI -0.69 to 0.97 ; informal, $g = 0.21$, 95% CI -0.62 to 1.04).

Help-Seeking Behavior

The interaction between condition and measurement occasion for help-seeking from formal sources was not significant for any condition compared with the control at postintervention (mental health literacy/destigmatization, OR 57.38, 95% CI 0.85–3868.09, $P = .06$; feedback, OR 5.15, 95% CI 0.04–637.04, $P = .51$; help-seeking list, OR 13.89, 95% CI 0.15–1263.93, $P = .25$) or at 3-month follow-up (mental health literacy/destigmatization, OR 3.48, 95% CI 0.10–122.32, $P = .49$; feedback, OR 9.45, 95% CI 0.18–507.02, $P = .27$; help-seeking list, OR 15.28, 95% CI 0.30–766.56, $P = .17$). Similarly, the overall interaction between condition and measurement occasion for help-seeking from informal sources was not significant for any condition compared with the control at postintervention (mental health literacy/destigmatization, OR 0.74, 95% CI 0.03–19.12, $P = .86$; feedback, OR 0.03, 95% CI 0.00–1.95, $P = .1$; help-seeking list, OR 1.07, 95% CI 0.03–35.22, $P = .97$) or at 3-month follow-up (mental health literacy/destigmatization, OR 0.21, 95% CI 0.01–7.79, $P = .4$; feedback, OR 0.01, 95% CI 0.00–1.21, $P = .06$; help-seeking list, OR 0.63, 95% CI 0.01–28.28, $P = .81$).

Secondary Intervention Effects

Table 4 presents the observed means and standard deviations for the secondary outcome variables, and the significance levels of the planned contrasts comparing change over time between each intervention condition and all other conditions.

Table 4. Observed means and standard deviations for the secondary outcome measures at pre- and postintervention for the trial conditions.^a

Measure	Measurement occasion					
	Preintervention		Postintervention		3-month follow-up	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Depression literacy (D-Lit)^b						
Mental health literacy/destigmatization	10	12.60 (3.63)	10	16.00 (3.50) ^c	7	14.71 (2.36) ^d
Feedback	9	12.67 (3.50)	11	12.73 (2.94)	7	13.14 (3.76)
Help-seeking list	9	10.67 (2.29)	12	10.92 (2.54)	9	11.33 (2.83)
Control	12	12.17 (3.33)	14	12.21 (4.73)	15	11.67 (5.34)
Anxiety literacy (A-Lit)^e						
Mental health literacy/destigmatization	10	9.00 (3.53)	10	13.70 (4.88) ^f	7	13.57 (4.16) ^g
Feedback	9	9.33 (4.15)	11	8.27 (3.23)	7	8.57 (3.95)
Help-seeking list	9	7.44 (3.84)	12	9.17 (3.19)	9	10.00 (3.64)
Control	12	8.08 (3.50)	14	9.57 (4.48)	15	10.13 (4.75)
Depression stigma (DSS)^h						
Mental health literacy/destigmatization	10	11.50 (2.76)	10	7.50 (4.95) ⁱ	7	7.71 (3.77) ^j
Feedback	9	8.44 (6.46)	11	9.82 (5.90)	7	8.57 (5.62)
Help-seeking list	9	11.11 (5.65)	12	7.58 (3.32)	9	9.56 (4.19) ^k
Control	12	9.50 (4.32)	14	8.93 (6.39)	15	8.13 (4.70)
Anxiety stigma (GASS)^l						
Mental health literacy/destigmatization	10	8.50 (3.27)	10	5.40 (4.77) ^m	7	3.71 (3.35) ⁿ
Feedback	8	4.38 (4.10)	11	8.18 (5.38)	7	6.57 (4.50)
Help-seeking list	9	8.67 (6.21)	12	6.00 (4.22)	9	5.78 (5.07)
Control	12	6.42 (6.07)	14	5.57 (4.36)	15	6.00 (5.07)

^a Significance tests refer to pattern of change.^b Depression Literacy questionnaire.^c Pre- to postintervention vs control, $P = .003$; vs help-seeking, $P = .004$.^d Preintervention to 3-month follow-up vs control, $P = .005$; vs help-seeking, $P = .04$.^e Anxiety Literacy questionnaire.^f Pre- to postintervention vs control, $P = .002$; vs help-seeking, $P = .003$; vs feedback, $P = .001$.^g Preintervention to 3-month follow-up vs control, $P = .001$; vs help-seeking, $P = .004$; vs feedback, $P < .001$.^h Depression Stigma Scale.ⁱ Pre- to postintervention vs control, $P = .01$; vs help-seeking, $P = .004$; vs feedback, $P = .003$.^j Preintervention to 3-month follow-up vs help-seeking, $P = .002$.^k Preintervention to 3-month follow-up vs control, $P = .04$.^l Generalised Anxiety Stigma Scale.^m Pre- to postintervention vs feedback, $P = .004$.ⁿ Preintervention to 3-month follow-up vs control, $P = .02$; vs help-seeking, $P = .03$; vs feedback, $P = .009$.

Depression Literacy

The overall interaction between condition and measurement occasion for depression literacy was significant ($F_{6,69.41} = 2.47$, $P = .03$). Planned contrasts demonstrated in the mental health literacy/destigmatization condition a greater increase in

depression literacy from pre- to postintervention than in the control and help-seeking list conditions (see Table 4), and approached significance for the feedback condition ($P = .05$). These effects were maintained at 3-month follow-up. There were no other significant effects. Between-group effect sizes at postintervention compared with control were as follows: mental

health literacy/destigmatization condition ($g = 0.90$, 95% CI 0.05–1.75), feedback condition ($g = 0.13$, 95% CI –0.66 to 0.92), and help-seeking list condition ($g = -0.34$, 95% CI –1.11 to 0.44). At 3-month follow-up, effect sizes compared with control for depression literacy were as follows: mental health literacy/destigmatization condition ($g = 0.66$, 95% CI –0.26 to 1.58), feedback condition ($g = 0.30$, 95% CI –0.60 to 1.20), and help-seeking list condition ($g = -0.07$, 95% CI –0.90 to 0.75). Given missing data at postintervention and 3-month follow-up and the lower precision of effect size confidence intervals than of planned contrasts, the effect size confidence intervals include zero even where the contrasts were significant. Patterns were similar for the literacy and stigma outcomes below.

Anxiety Literacy

The overall interaction between condition and measurement occasion for anxiety literacy was significant ($F_{6,67.51} = 3.99$, $P = .002$). Planned contrasts demonstrated in the mental health literacy/destigmatization condition a greater increase in anxiety literacy than in all other conditions at postintervention, and these effects were maintained at 3-month follow-up (see Table 4). There were no other significant effects. Between-group effect sizes at postintervention relative to control were as follows: mental health literacy/destigmatization condition ($g = 0.90$, 95% CI 0.05–1.75), feedback condition ($g = -0.33$, 95% CI –1.12 to 0.47), and help-seeking list condition ($g = -0.10$, 95% CI –0.87 to 0.67). At 3-month follow-up, effect sizes compared with control for anxiety literacy were as follows: mental health literacy/destigmatization condition ($g = 0.76$, 95% CI –0.17 to 1.68), feedback condition ($g = -0.35$, 95% CI –1.25 to 0.56), and help-seeking list condition ($g = -0.03$, 95% CI –0.74 to 0.69).

Depression Stigma

The overall interaction between condition and measurement occasion for depression personal stigma was statistically significant ($F_{6,62.22} = 3.20$, $P = .008$). Planned contrasts demonstrated in the mental health literacy/destigmatization condition a greater decrease in depression stigma than in all other conditions from pre- to postintervention (see Table 4). At 3-month follow-up the effects for the mental health literacy/destigmatization condition were no longer superior for depression stigma compared with any condition, except for the help-seeking list condition, which was associated with greater stigma. There were no other significant effects. Between-group effect sizes at postintervention relative to control were as follows: mental health literacy/destigmatization condition ($g = 0.25$, 95% CI –0.57 to 1.06), feedback condition ($g = -0.15$, 95% CI –0.94 to 0.65), and help-seeking list condition ($g = 0.26$, 95% CI –0.51 to 1.04). At 3-month follow-up, effect sizes compared with control for depression stigma were as follows: mental health literacy/destigmatization condition ($g = 0.10$, 95% CI –0.80 to 0.99), feedback condition ($g = -0.09$, 95% CI –0.99 to 0.81), and help-seeking list condition ($g = -0.32$, 95% CI –1.15 to 0.51).

Anxiety Stigma

The overall interaction between condition and measurement occasion for anxiety stigma was significant ($F_{6,65.37} = 2.27$, $P =$

.047). Planned contrasts demonstrated in the mental health literacy/destigmatization condition a greater decrease in anxiety stigma from pre- to postintervention than in the feedback condition only (see Table 4). However, at 3-month follow-up the reduction in anxiety stigma for the mental health literacy/destigmatization condition was significant compared with all conditions. There were no other significant effects. Between-group effect sizes at postintervention relative to control were as follows: mental health literacy/destigmatization condition ($g = 0.04$, 95% CI –0.77 to 0.85), feedback condition ($g = -0.54$, 95% CI –1.35 to 0.26), and help-seeking list condition ($g = -0.10$, 95% CI –0.87 to 0.67). At 3-month follow-up, effect sizes compared with the control for anxiety stigma were as follows: mental health literacy/destigmatization condition ($g = 0.50$, 95% CI –0.41 to 1.41), feedback condition ($g = 0.12$, 95% CI –1.02 to 0.78), and help-seeking list condition ($g = 0.04$, 95% CI –0.78 to 0.87).

Discussion

Principal Findings

The present study examined the effects of three brief Internet-based interventions designed to increase help-seeking attitudes, intentions, and behavior in elite athletes. None of the interventions were efficacious in improving mental health help-seeking attitudes, intentions, or behavior relative to the control condition. By contrast, several previous RCTs have demonstrated improvements in help-seeking attitudes for mental health conditions following brief interventions [12–14].

It is likely that the negative finding for help-seeking attitudes in the current study was due to low power, given that it was not feasible to recruit the target sample size. This is supported by our finding that at postintervention the effect size for help-seeking attitudes for the mental health literacy/destigmatization condition ($g = 0.28$) is comparable with those of previous studies, which reported significant intervention effects (between-group effect sizes: d range of 0.12–0.34) [12–14]. There was also a trend for an increase in formal help-seeking behavior in the mental health literacy/destigmatization condition compared with the control group at postintervention.

It may be relevant that the athletes in the present study had more positive attitudes at preintervention overall (mean score 20.85) than members of a general population sample of a similar age (mean 17.10) [14]. This finding contrasts with the results of a previous research study, which reported that US college athletes had less positive attitudes than their nonathlete peers toward help-seeking for mental health problems from counselors [8]. The higher positive attitudes among athletes in the current sample may have been due to the high proportion of female participants, as women have been found to have more positive attitudes to help-seeking than men [48]. Consistent with this, the female participants in the present sample had more positive help-seeking attitudes than the male participants at preintervention, although this difference was not significant. Further, the significantly lower levels of participation by men in the present study suggest that they may be more difficult to engage with this type of intervention.

The present study found no effect of the interventions on overall help-seeking *intentions*, or on formal or informal help-seeking intentions, for any of the intervention conditions. The negative finding for intentions is consistent with a previous RCT of help-seeking intentions, which used the same measurement scale [17]. It differs somewhat from that of Costin et al [17], who found a significant pre- to postintervention change for help-seeking intentions from formal sources in both the intervention and control conditions. However, again, the present study sample of athletes had relatively high intentions (mean 4.22 out of 7.00) to seek help from formal sources at preintervention. Thus, it is possible that the athlete sample in the current study had less room to improve than participants in the Costin et al study (mean 2.61). Overall, it is encouraging that this sample of athletes had preexisting high levels of intentions to seek help. It suggests that it is important to ensure that those from whom they are likely to seek help are primed and able to deliver evidence-based interventions. It is unclear to what extent the high levels of help-seeking intention were attributable to the relatively high proportion of women in the current study [48]. It was not possible to investigate gender effects in the current study due to the small sample sizes achieved. However, further research should be undertaken to determine whether these interventions exert a differential effect on male and female athletes.

Few RCTs of interventions have investigated or demonstrated an effect on mental health help-seeking behavior. One of the strengths of the design of the present study was the measurement of help-seeking behavior at preintervention to allow comparison between measurement points. In the current study, the change in proportions from pre- to postintervention in those seeking formal help approached significance for the mental health literacy/destigmatization condition. While the sample size is small, these findings suggest that the mental health literacy/destigmatization intervention may be a promising tool for this population, especially given that the trial examined a universal sample not selected for their level of symptomatology. The only previous report of increased help-seeking behavior, by Christensen et al [16], was based on an indicated sample of members of the community with high levels of depressive symptoms.

The interventions did not significantly increase help-seeking behavior from informal sources. However, compared with professional sources, help-seeking behavior from informal sources was high at pre- and postintervention. The fact that athletes were already high seekers of help from informal sources allowed limited room for an increase in this behavior. Moreover, the focus of the interventions was on promoting formal help-seeking.

The mental health literacy/destigmatization condition led to improved depression and anxiety literacy relative to control, an effect that was maintained at 3-month follow-up. It has been argued that increasing mental health literacy in the community is a key means for facilitating help-seeking behavior by individuals and for enabling them to assist others in need [49]. In accordance with this, Han et al [15] concluded that providing psychoeducation (eg, mental health literacy) concerning biological attribution information for depression is an effective

method for increasing help-seeking willingness. It is encouraging that the mental health literacy/destigmatization condition in the present study successfully increased knowledge of both depression and anxiety in a brief Internet-based format.

The present study also suggested that the mental health/destigmatization intervention shows promise for decreasing stigma for depression and anxiety. The findings were not as robust as for literacy, with the effect disappearing at 3 months for depression stigma and evident only at 3 month follow up for anxiety stigma. It is possible that stigmatizing beliefs are more difficult to change, given the present study's stronger results for literacy (g range of 0.66–0.90), than stigma (g range of 0.04–0.50), although the effects for literacy may be attributable to floor effects. Nevertheless, this preliminary suggestion of a positive effect for stigma is encouraging given that stigma has been implicated as a barrier to help-seeking in young people [11], particularly for depression in adults [50,51].

Strengths and Limitations

To our knowledge, this is the first study to test the feasibility of targeting help-seeking attitudes, intentions, and behavior in an Internet-based format with elite athletes. Strengths of the study design were that it was universal, targeting not only those in current need of help, but also those who may need to seek help in the future. The focus on online delivery maximized the potential of the intervention to reach a large number and broad range of athletes cost effectively [19] and regardless of geographic location [21]. A further strength of the study was that we were able to track and document a moderately high level of adherence to the intervention. Including active intervention exercises in future interventions may further enhance our ability to track engagement. Finally, the inclusion of three different types of help-seeking interventions provided some preliminary data on the relative effect of each of these methods on help-seeking, stigma, and mental health literacy.

The low achieved sample was the primary limitation of the present study; this decreases the power of the study (increasing chance of type II error). The differing methods of recruitment may have also compromised the external validity of the trial. In addition, the recruitment methods may have resulted in selection bias, in that those who participated in the study may have been generally more interested in the research topic of mental health. As can be seen in Figure 1, athletes were willing to provide a response to the stage 1 survey, whereas only 59 (7.7%) submitted a pre- or a postintervention survey for the RCT. The fact that respondents who agreed to participate in the intervention were not a representative sample of those who completed the initial survey raises the possibility that those who agreed to participate already had positive attitudes toward help-seeking. This may have militated against finding positive effects. The potential for such bias is unlikely to be restricted to the current study, having been noted previously [17]; rather it is a challenge for all studies of help-seeking. Finally, a large majority of the participants were female, which also limits the generalizability of the results. Alternative recruitment strategies targeting low help-seekers, particularly men, are required.

The low recruitment rates may have arisen in part from the inclusion of the lengthy stage 1 survey prior to the invitation to

participate in the RCT, with the stage 1 survey inadvertently deterring participation in the RCT. This is an important consideration for future research, especially if the pool of potential participants is small and they have limited time due to demanding schedules as was the case in the present study. In addition, organizing participation by providing computer facilities at the club site markedly increased response rates to the stage 1 survey. The extension of such facilities to the intervention phase might have increased recruitment rates to the trial. Analyses of the factors associated with agreement to participate in the intervention component revealed that certain groups are less amenable to recruitment; there may be value in examining these predictors further with a view to tailoring recruitment strategies to this group. A final limitation is that the changes in the stigma scores might reflect a response bias, whereby the participants in the mental health literacy/destigmatization condition learned how to respond to the stigma items as expected but did not undergo a genuine change in attitudes to depression and anxiety.

Conclusions

Online interventions that can facilitate mental health help-seeking in targeted populations such as elite athletes have the potential to play an important role in decreasing the prevalence of depression and anxiety disorders in these populations. This exploratory study found that a mental health literacy/destigmatization intervention increased athletes' knowledge of depression and anxiety disorders, and showed some evidence of reducing their depression and anxiety stigma. It also showed a nonsignificant trend toward improving help-seeking behavior among athletes. Further research is required to evaluate the effect of the mental health literacy/destigmatization intervention employing a sample size that is sufficient to detect a moderate effect size. Further work should be undertaken with athletes and in a range of other settings including schools and workplaces, and including more active components in the intervention to enhance engagement should be considered.

Acknowledgments

We wish to thank Matti Clements, Professor Peter Fricker, and the Australian Institute of Sport (AIS) for assisting with the recruitment of respondents; Michael Martin, Lyndel Mayoh, and Ruth Anderson for their contribution to the early design of the study; and Sally Hughson for the final-stage supervision of the project. The study was supported by a grant from the AIS. AG is supported by a joint scholarship from the AIS, the Brain and Mind Research Institute, Orygen, and The Australian National University. KG is supported by NHMRC Senior Research Fellowship No. 525413. HC is supported by NHMRC Senior Principal Research Fellowship No. 525411. AC is supported by NHMRC Early Career Fellowship 1013199. PB is supported by NHMRC Capacity Building Grant 418020.

Conflicts of Interest

The present study authors developed the TEAMS online interventions but derived no personal or financial benefit from their operation.

References

1. Wittchen HU. Generalized anxiety disorder: prevalence, burden, and cost to society. *Depress Anxiety* 2002;16(4):162-171. [doi: [10.1002/da.10065](https://doi.org/10.1002/da.10065)] [Medline: [12497648](https://pubmed.ncbi.nlm.nih.gov/12497648/)]
2. Andrews G, Issakidis C, Carter G. Shortfall in mental health service utilisation. *Br J Psychiatry* 2001 Nov;179:417-425 [FREE Full text] [Medline: [11689399](https://pubmed.ncbi.nlm.nih.gov/11689399/)]
3. Anonymous. Australian Bureau of Statistics. 2008 Oct 23. 4326.0: National Survey of Mental Health and Wellbeing: Summary of Results, 2007 URL: <http://www.abs.gov.au/ausstats/abs@.nsf/mf/4326.0> [accessed 2012-04-25] [WebCite Cache ID 67C7O3ekM]
4. King M, Nazareth I, Levy G, Walker C, Morris R, Weich S, et al. Prevalence of common mental disorders in general practice attendees across Europe. *Br J Psychiatry* 2008 May;192(5):362-367 [FREE Full text] [doi: [10.1192/bjp.bp.107.039966](https://doi.org/10.1192/bjp.bp.107.039966)] [Medline: [18450661](https://pubmed.ncbi.nlm.nih.gov/18450661/)]
5. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun;62(6):617-627 [FREE Full text] [doi: [10.1001/archpsyc.62.6.617](https://doi.org/10.1001/archpsyc.62.6.617)] [Medline: [15939839](https://pubmed.ncbi.nlm.nih.gov/15939839/)]
6. Rickwood DJ, Deane FP, Wilson CJ. When and how do young people seek professional help for mental health problems? *Med J Aust* 2007 Oct 1;187(7 Suppl):S35-S39 [FREE Full text] [Medline: [17908023](https://pubmed.ncbi.nlm.nih.gov/17908023/)]
7. Yang J, Peek-Asa C, Corlette JD, Cheng G, Foster DT, Albright J. Prevalence of and risk factors associated with symptoms of depression in competitive collegiate student athletes. *Clin J Sport Med* 2007 Nov;17(6):481-487. [doi: [10.1097/JSM.0b013e31815aed6b](https://doi.org/10.1097/JSM.0b013e31815aed6b)] [Medline: [17993792](https://pubmed.ncbi.nlm.nih.gov/17993792/)]
8. Watson J. College student-athletes' attitudes toward help-seeking behavior and expectations of counseling services. *J Coll Stud Dev* 2005;46(4):442-449.
9. Komiti A, Judd F, Jackson H. The influence of stigma and attitudes on seeking help from a GP for mental health problems: a rural context. *Soc Psychiatry Psychiatr Epidemiol* 2006 Sep;41(9):738-745. [doi: [10.1007/s00127-006-0089-4](https://doi.org/10.1007/s00127-006-0089-4)] [Medline: [16819564](https://pubmed.ncbi.nlm.nih.gov/16819564/)]

10. Kushner M, Sher K. The relation of treatment fearfulness and psychological service utilization: an overview. *Prof Psychol Res Pr* 1991 Sep;22(3):196-203. [doi: [10.1037/0735-7028.22.3.196](https://doi.org/10.1037/0735-7028.22.3.196)]
11. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry* 2010;10:113 [FREE Full text] [doi: [10.1186/1471-244X-10-113](https://doi.org/10.1186/1471-244X-10-113)] [Medline: [21192795](https://pubmed.ncbi.nlm.nih.gov/21192795/)]
12. Buckley GI, Malouff JM. Using modeling and vicarious reinforcement to produce more positive attitudes toward mental health treatment. *J Psychol* 2005 May;139(3):197-209. [Medline: [15945515](https://pubmed.ncbi.nlm.nih.gov/15945515/)]
13. Donohue B, Dickens Y, Lancer K, Covassin T, Hash A, Miller A, et al. Improving athletes' perspectives of sport psychology consultation: a controlled evaluation of two interview methods. *Behav Modif* 2004 Mar;28(2):182-193. [doi: [10.1177/0145445503259399](https://doi.org/10.1177/0145445503259399)] [Medline: [14997947](https://pubmed.ncbi.nlm.nih.gov/14997947/)]
14. Sharp W, Hargrove D, Johnson L, Deal W. Help-seeking and mental health education: an evaluation of a classroom-based strategy to modify help-seeking for mental health problems. *J Coll Stud Dev* 2006;47(4):419-438.
15. Han DY, Chen SH, Hwang KK, Wei HL. Effects of psychoeducation for depression on help-seeking willingness: biological attribution versus destigmatization. *Psychiatry Clin Neurosci* 2006 Dec;60(6):662-668. [doi: [10.1111/j.1440-1819.2006.01580.x](https://doi.org/10.1111/j.1440-1819.2006.01580.x)] [Medline: [17109699](https://pubmed.ncbi.nlm.nih.gov/17109699/)]
16. Christensen H, Leach LS, Barney L, Mackinnon AJ, Griffiths KM. The effect of web based depression interventions on self reported help seeking: randomised controlled trial [ISRCTN77824516]. *BMC Psychiatry* 2006;6:13 [FREE Full text] [doi: [10.1186/1471-244X-6-13](https://doi.org/10.1186/1471-244X-6-13)] [Medline: [16595018](https://pubmed.ncbi.nlm.nih.gov/16595018/)]
17. Costin DL, Mackinnon AJ, Griffiths KM, Batterham PJ, Bennett AJ, Bennett K, et al. Health e-cards as a means of encouraging help seeking for depression among young adults: randomized controlled trial. *J Med Internet Res* 2009;11(4):e42 [FREE Full text] [doi: [10.2196/jmir.1294](https://doi.org/10.2196/jmir.1294)] [Medline: [19850549](https://pubmed.ncbi.nlm.nih.gov/19850549/)]
18. Rickwood D, Deane F, Wilson C, Ciarrochi J. Young people's help-seeking for mental health problems. *Adv Ment Health* 2005;4(3):218-251. [doi: [10.5172/jamh.4.3.218](https://doi.org/10.5172/jamh.4.3.218)]
19. Gerhards SA, de Graaf LE, Jacobs LE, Severens JL, Huibers MJ, Arntz A, et al. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. *Br J Psychiatry* 2010 Apr;196(4):310-318 [FREE Full text] [doi: [10.1192/bjp.bp.109.065748](https://doi.org/10.1192/bjp.bp.109.065748)] [Medline: [20357309](https://pubmed.ncbi.nlm.nih.gov/20357309/)]
20. Mihalopoulos C, Kiropoulos L, Shih ST, Gunn J, Blaschki G, Meadows G. Exploratory economic analyses of two primary care mental health projects: implications for sustainability. *Med J Aust* 2005 Nov 21;183(10 Suppl):S73-S76 [FREE Full text] [Medline: [16296957](https://pubmed.ncbi.nlm.nih.gov/16296957/)]
21. Napolitano MA, Marcus BH. Targeting and tailoring physical activity information using print and information technologies. *Exerc Sport Sci Rev* 2002 Jul;30(3):122-128. [Medline: [12150571](https://pubmed.ncbi.nlm.nih.gov/12150571/)]
22. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. *J Med Internet Res* 2010;12(5):e60 [FREE Full text] [doi: [10.2196/jmir.1463](https://doi.org/10.2196/jmir.1463)] [Medline: [21169162](https://pubmed.ncbi.nlm.nih.gov/21169162/)]
23. Fridrici M, Lohaus A, Glass C. Effects of incentives in web-based prevention for adolescents: Results of an exploratory field study. *Psychol Health* 2009 Jul;24(6):663-675. [doi: [10.1080/08870440802521102](https://doi.org/10.1080/08870440802521102)] [Medline: [20205019](https://pubmed.ncbi.nlm.nih.gov/20205019/)]
24. Christensen H, Griffiths K, Groves C. MoodGYM Training Program: Clinician's Manual. Canberra: Centre for Mental Health Research; 2004.
25. Jorm AF, Dear KB, Rodgers B, Christensen H. Interaction between mother's and father's affection as a risk factor for anxiety and depression symptoms--evidence for increased risk in adults who rate their father as having been more affectionate than their mother. *Soc Psychiatry Psychiatr Epidemiol* 2003 Apr;38(4):173-179. [doi: [10.1007/s00127-003-0620-9](https://doi.org/10.1007/s00127-003-0620-9)] [Medline: [12664227](https://pubmed.ncbi.nlm.nih.gov/12664227/)]
26. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. *Health Educ Res* 2005 Oct;20(5):527-539 [FREE Full text] [doi: [10.1093/her/cyh014](https://doi.org/10.1093/her/cyh014)] [Medline: [15701665](https://pubmed.ncbi.nlm.nih.gov/15701665/)]
27. Bewick BM, West R, Gill J, O'May F, Mulhern B, Barkham M, et al. Providing web-based feedback and social norms information to reduce student alcohol intake: a multisite investigation. *J Med Internet Res* 2010;12(5):e59 [FREE Full text] [doi: [10.2196/jmir.1461](https://doi.org/10.2196/jmir.1461)] [Medline: [21169171](https://pubmed.ncbi.nlm.nih.gov/21169171/)]
28. Burgess SW, Sly PD, Devadason SG. Providing feedback on adherence increases use of preventive medication by asthmatic children. *J Asthma* 2010 Mar;47(2):198-201. [doi: [10.3109/02770900903483840](https://doi.org/10.3109/02770900903483840)] [Medline: [20170329](https://pubmed.ncbi.nlm.nih.gov/20170329/)]
29. Rinfret S, Lussier MT, Peirce A, Duhamel F, Cossette S, Lalonde L, LOYAL Study Investigators. The impact of a multidisciplinary information technology-supported program on blood pressure control in primary care. *Circ Cardiovasc Qual Outcomes* 2009 May;2(3):170-177 [FREE Full text] [doi: [10.1161/CIRCOUTCOMES.108.823765](https://doi.org/10.1161/CIRCOUTCOMES.108.823765)] [Medline: [20031834](https://pubmed.ncbi.nlm.nih.gov/20031834/)]
30. Bien TH, Miller WR, Tonigan JS. Brief interventions for alcohol problems: a review. *Addiction* 1993 Mar;88(3):315-335. [Medline: [8461850](https://pubmed.ncbi.nlm.nih.gov/8461850/)]
31. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002 Aug;32(6):959-976. [Medline: [12214795](https://pubmed.ncbi.nlm.nih.gov/12214795/)]
32. Wilson C, Deane F, Ciarrochi J, Rickwood D. Measuring help-seeking intentions: properties of the General Help-Seeking Questionnaire. *Can J Couns* 2005;39(1):15-28.
33. Fischer EH, Farina A. Attitudes toward seeking professional psychological help: a shortened form and considerations for research. *J Coll Stud Dev* 1995;36(4):368-373.

34. Elhai JD, Schweinle W, Anderson SM. Reliability and validity of the Attitudes Toward Seeking Professional Psychological Help Scale-Short Form. *Psychiatry Res* 2008 Jun 30;159(3):320-329. [doi: [10.1016/j.psychres.2007.04.020](https://doi.org/10.1016/j.psychres.2007.04.020)] [Medline: [18433879](https://pubmed.ncbi.nlm.nih.gov/18433879/)]
35. Hackler A. Attitudes toward seeking professional help for an eating disorder: the role of stigma and anticipated outcomes. *J Couns Dev* 2010;88(4):424-431.
36. Chang HW. Help-seeking for stressful events among Chinese college students in Taiwan: roles of gender, prior history of counseling, and help-seeking attitudes. *J Coll Stud Dev* 2008;49(1):41-51.
37. Deane FP, Wilson CJ, Ciarrochi J. Suicidal ideation and help-negation: Not just hopelessness or prior help. *J Clin Psychol* 2001 Jul;57(7):901-914. [Medline: [11406803](https://pubmed.ncbi.nlm.nih.gov/11406803/)]
38. Deane FP, Wilson CJ. University of Wollongong. 2007 Oct 2. Considerations for Specifying Problem-Types, Help-Sources and Scoring the General Help-seeking Questionnaire (GHSQ) URL: <http://www.uow.edu.au/content/groups/public/@web/@health/documents/doc/uow039041.pdf> [accessed 2009-06-09] [WebCite Cache ID 5hOaNux4b]
39. Rickwood D, Cavanagh S, Curtis L, Sakrouge R. Educating young people about mental health and mental illness: evaluating a school-based programme. *Int J Ment Health Promot* 2004;6(4):23-32.
40. Rickwood DJ, Braithwaite VA. Social-psychological factors affecting help-seeking for emotional problems. *Soc Sci Med* 1994 Aug;39(4):563-572. [Medline: [7973856](https://pubmed.ncbi.nlm.nih.gov/7973856/)]
41. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. *Br J Psychiatry* 2004 Oct;185:342-349 [FREE Full text] [doi: [10.1192/bjp.185.4.342](https://doi.org/10.1192/bjp.185.4.342)] [Medline: [15458995](https://pubmed.ncbi.nlm.nih.gov/15458995/)]
42. Kiropoulos LA, Griffiths KM, Blashki G. Effects of a multilingual information website intervention on the levels of depression literacy and depression-related stigma in Greek-born and Italian-born immigrants living in Australia: a randomized controlled trial. *J Med Internet Res* 2011;13(2):e34 [FREE Full text] [doi: [10.2196/jmir.1527](https://doi.org/10.2196/jmir.1527)] [Medline: [21504872](https://pubmed.ncbi.nlm.nih.gov/21504872/)]
43. Griffiths KM, Batterham PJ, Barney L, Parsons A. The Generalised Anxiety Stigma Scale (GASS): psychometric properties in a community sample. *BMC Psychiatry* 2011;11:184 [FREE Full text] [doi: [10.1186/1471-244X-11-184](https://doi.org/10.1186/1471-244X-11-184)] [Medline: [22108099](https://pubmed.ncbi.nlm.nih.gov/22108099/)]
44. Griffiths KM, Christensen H, Jorm AF. Predictors of depression stigma. *BMC Psychiatry* 2008;8:25 [FREE Full text] [doi: [10.1186/1471-244X-8-25](https://doi.org/10.1186/1471-244X-8-25)] [Medline: [18423003](https://pubmed.ncbi.nlm.nih.gov/18423003/)]
45. Gueorgieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Arch Gen Psychiatry* 2004 Mar;61(3):310-317. [doi: [10.1001/archpsyc.61.3.310](https://doi.org/10.1001/archpsyc.61.3.310)] [Medline: [14993119](https://pubmed.ncbi.nlm.nih.gov/14993119/)]
46. Hamer RM, Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *Am J Psychiatry* 2009 Jun;166(6):639-641. [doi: [10.1176/appi.ajp.2009.09040458](https://doi.org/10.1176/appi.ajp.2009.09040458)] [Medline: [19487398](https://pubmed.ncbi.nlm.nih.gov/19487398/)]
47. Hedges L, Olkin I. *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press; 1985.
48. Mackenzie CS, Gekoski WL, Knox VJ. Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. *Aging Ment Health* 2006 Nov;10(6):574-582. [doi: [10.1080/13607860600641200](https://doi.org/10.1080/13607860600641200)] [Medline: [17050086](https://pubmed.ncbi.nlm.nih.gov/17050086/)]
49. Kelly CM, Jorm AF, Wright A. Improving mental health literacy as a strategy to facilitate early intervention for mental disorders. *Med J Aust* 2007 Oct 1;187(7 Suppl):S26-S30 [FREE Full text] [Medline: [17908021](https://pubmed.ncbi.nlm.nih.gov/17908021/)]
50. Barney LJ, Griffiths KM, Christensen H, Jorm AF. Exploring the nature of stigmatising beliefs about depression and help-seeking: implications for reducing stigma. *BMC Public Health* 2009;9:61 [FREE Full text] [doi: [10.1186/1471-2458-9-61](https://doi.org/10.1186/1471-2458-9-61)] [Medline: [19228435](https://pubmed.ncbi.nlm.nih.gov/19228435/)]
51. Barney LJ, Griffiths KM, Jorm AF, Christensen H. Stigma about depression and its impact on help-seeking intentions. *Aust N Z J Psychiatry* 2006 Jan;40(1):51-54 [FREE Full text] [doi: [10.1111/j.1440-1614.2006.01741.x](https://doi.org/10.1111/j.1440-1614.2006.01741.x)] [Medline: [16403038](https://pubmed.ncbi.nlm.nih.gov/16403038/)]

Abbreviations

AHSQ: Actual Help-Seeking Questionnaire
AIS: Australian Institute of Sport
A-Lit: Anxiety Literacy questionnaire
ATSPPH-SF: Attitudes Toward Seeking Professional Psychological Help-Short Form
CI: confidence interval
D-Lit: Depression Literacy questionnaire
DSS: Depression Stigma Scale
GASS: Generalised Anxiety Stigma Scale
GHSQ: General Help-Seeking Questionnaire
K10: 10-item Kessler Psychological Distress Scale
OR: odds ratio
RCT: randomized controlled trial
TEAMS: The Elite Athlete Mental Health Strategy

Edited by F Thorndike; submitted 08.06.11; peer-reviewed by D Broshek, S Reid; comments to author 16.08.11; revised version received 04.11.11; accepted 06.02.12; published 29.06.12.

Please cite as:

Gulliver A, Griffiths KM, Christensen H, Mackinnon A, Cleave AL, Parsons A, Bennett K, Batterham PJ, Stanimirovic R
Internet-Based Interventions to Promote Mental Health Help-Seeking in Elite Athletes: An Exploratory Randomized Controlled Trial
J Med Internet Res 2012;14(3):e69

URL: <http://www.jmir.org/2012/3/e69/>

doi: [10.2196/jmir.1864](https://doi.org/10.2196/jmir.1864)

PMID: [22743352](https://pubmed.ncbi.nlm.nih.gov/22743352/)

©Amelia Gulliver, Kathleen M Griffiths, Helen Christensen, Andrew Mackinnon, Alison L Cleave, Alison Parsons, Kylie Bennett, Philip J Batterham, Rosanna Stanimirovic. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Letter to the Editor

From eHealth Technologies to Interventions

Rik Crutzen¹, PhD

Maastricht University/CAPHRI, Maastricht, Netherlands

Corresponding Author:

Rik Crutzen, PhD

Maastricht University/CAPHRI

P.O. Box 616

Maastricht, 6200 MD

Netherlands

Phone: 31 433882828

Fax: 31 433671032

Email: rik.crutzen@maastrichtuniversity.nl

(*J Med Internet Res* 2012;14(3):e93) doi:[10.2196/jmir.2050](https://doi.org/10.2196/jmir.2050)

KEYWORDS

eHealth technologies; interventions; Intervention Mapping

Van Gemert-Pijnen and colleagues recently presented an impressive viewpoint paper describing a holistic framework for the development of eHealth technologies, based on a review of existing eHealth frameworks [1]. As this review was limited to eHealth frameworks, the resulting holistic framework is focused on creating technologies instead of interventions. In my and others' opinion, however, "interactive technologies are a tool, not a panacea" [2]. Therefore, it is useful – in *addition* to the holistic framework for the development of eHealth technologies – to look at intervention development approaches in general. Intervention Mapping, for example, focuses on the planning and development of effective interventions, not limited to eHealth technologies [3]. Comparable to the holistic framework, this is an iterative approach that takes into account social dynamics by incorporating stakeholders in a so-called linkage group. This linkage group consists of end users (e.g., patients) as well as intermediaries (e.g., nurse practitioners) and decision makers (e.g., department head, policy makers). It is striking to see the overlap between the activities mentioned in the holistic framework and the steps to be taken within Intervention Mapping (Table 1). Besides the overlap, Table 1 also indicates the differences in terminology used throughout the literature. Contextual inquiry, for example, entails "information gathering from the intended users and the environment in which the technology will be implemented" [1]. This is comparable to a key component of a needs assessment which encompasses "an effort to understand the character of the community, its members, and its strengths" [3]. During value specification, values are determined and ranked "based on the importance of finding solutions for the identified problems" [1]. One of the key tasks in the second step of Intervention Mapping is to select important and changeable determinants of behavioral and environmental outcomes [3].

More important than differences in terminology, however, are substantive differences between the holistic framework and

Intervention Mapping. The two most important differences refer to the impact and uptake of eHealth technologies (both concepts mentioned in the title of the viewpoint paper [1]).

First, splitting up the design activity in two successive steps in Intervention Mapping is not merely a case of semantics or being more detailed. Active ingredients in interventions need to be based on theory-based intervention methods. These methods are to be translated into practical applications. Subsequently, these practical applications are incorporated while producing intervention programs. This results, finally, in intervention programs targeted at important and changeable determinants of behavioral and environmental outcomes, thereby increasing the likelihood of effective interventions (i.e., increasing their impact). Furthermore, it is insightful to look at active ingredients in interventions to unravel what works [4, 5], which can be facilitated by specifying methods and applications beforehand.

Second, with regard to uptake, the idea of Van Gemert-Pijnen and colleagues that long-term implementation requires business modeling is accepted with open arms. Although the role of cost-effectiveness analysis in health and medicine has been acknowledged for a while [6], business modeling is the next step that needs to be taken. Concepts and techniques from business modeling help to identify critical factors for the implementation, as Van Limburg and colleagues elaborated on in their companion paper [7].

Based on these differences, this letter should be seen as a call for integration: to treat interventions as technologies or even products that need to be marketed and require business modeling, as well as to focus on the content of interventions by means of producing intervention programs based on theory-based intervention methods that are translated into practical applications. In this way, eHealth interventions can effectively change health risk behaviors and their determinants

[8] and their long-term public health potential can be actualized [9].

Table 1. Activities in the holistic framework and steps within Intervention Mapping

Holistic framework	Intervention Mapping
Contextual inquiry	Needs assessment
Value specification	Matrices of change objectives
Design	Theory-based intervention methods and practical applications
	Intervention program
Operationalization	Adoption and implementation
Summative evaluation	Evaluation

Acknowledgments

This work was supported by the Innovational Research Incentives Scheme Veni from NWO-MaGW (Netherlands Organisation for Scientific Research - Division for the Social Sciences).

Conflicts of Interest

None declared.

References

1. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](https://doi.org/10.2196/jmir.1672)] [Medline: [22155738](https://pubmed.ncbi.nlm.nih.gov/22155738/)]
2. Bandura A. Health promotion by social cognitive means. *Health Educ Behav* 2004 Apr;31(2):143-164. [doi: [10.1177/1090198104263660](https://doi.org/10.1177/1090198104263660)] [Medline: [15090118](https://pubmed.ncbi.nlm.nih.gov/15090118/)]
3. Bartholomew LK, Parcel GS, Kok G, Gottlieb NH, Fernández ME. *Planning Health Promotion Programs: An Intervention Mapping Approach*. San Francisco, CA: Jossey-Bass; 2011.
4. Michie S. What works and how? Designing more effective interventions needs answers to both questions. *Addiction* 2008;103:886-887. [doi: [10.1111/j.1360-0443.2007.02112.x](https://doi.org/10.1111/j.1360-0443.2007.02112.x)]
5. Crutzen R. Is it time to start focusing on the content of computer-delivered interventions? *Addiction* 2011 Oct;106(10):1871. [doi: [10.1111/j.1360-0443.2011.03539.x](https://doi.org/10.1111/j.1360-0443.2011.03539.x)] [Medline: [21793973](https://pubmed.ncbi.nlm.nih.gov/21793973/)]
6. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996 Oct 9;276(14):1172-1177. [Medline: [8827972](https://pubmed.ncbi.nlm.nih.gov/8827972/)]
7. van Limburg M, van Gemert-Pijnen JE, Nijland N, Ossebaard HC, Hendrix RM, Seydel ER. Why business modeling is crucial in the development of eHealth technologies. *J Med Internet Res* 2011;13(4):e124 [FREE Full text] [doi: [10.2196/jmir.1674](https://doi.org/10.2196/jmir.1674)] [Medline: [22204896](https://pubmed.ncbi.nlm.nih.gov/22204896/)]
8. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. *Prev Med* 2008 Jul;47(1):3-16. [doi: [10.1016/j.ypmed.2008.02.014](https://doi.org/10.1016/j.ypmed.2008.02.014)] [Medline: [18403003](https://pubmed.ncbi.nlm.nih.gov/18403003/)]
9. 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](https://doi.org/10.1146/annurev.publhealth.031308.100235)] [Medline: [19296777](https://pubmed.ncbi.nlm.nih.gov/19296777/)]

Edited by H Potts; submitted 11.01.12; this is a non-peer-reviewed article; accepted 09.03.12; published 27.06.12.

Please cite as:

Crutzen R

From eHealth Technologies to Interventions

J Med Internet Res 2012;14(3):e93

URL: <http://www.jmir.org/2012/3/e93/>

doi: [10.2196/jmir.2050](https://doi.org/10.2196/jmir.2050)

PMID: [22738742](https://pubmed.ncbi.nlm.nih.gov/22738742/)

©Rik Crutzen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Letter to the Editor

Consistently Increasing Numbers of Online Ratings of Healthcare in England

Felix Greaves¹, BM, BCh, MPH, MBA; Christopher Millett¹, PhD

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

Corresponding Author:

Felix Greaves, BM, BCh, MPH, MBA

Department of Primary Care and Public Health

Imperial College London

Reynolds Building, Charing Cross Campus

London, W6 8RP

United Kingdom

Phone: 44 7866 551172

Fax: 44 207 5940584

Email: felix.greaves08@imperial.ac.uk

(*J Med Internet Res* 2012;14(3):e94) doi:[10.2196/jmir.2157](https://doi.org/10.2196/jmir.2157)

KEYWORDS

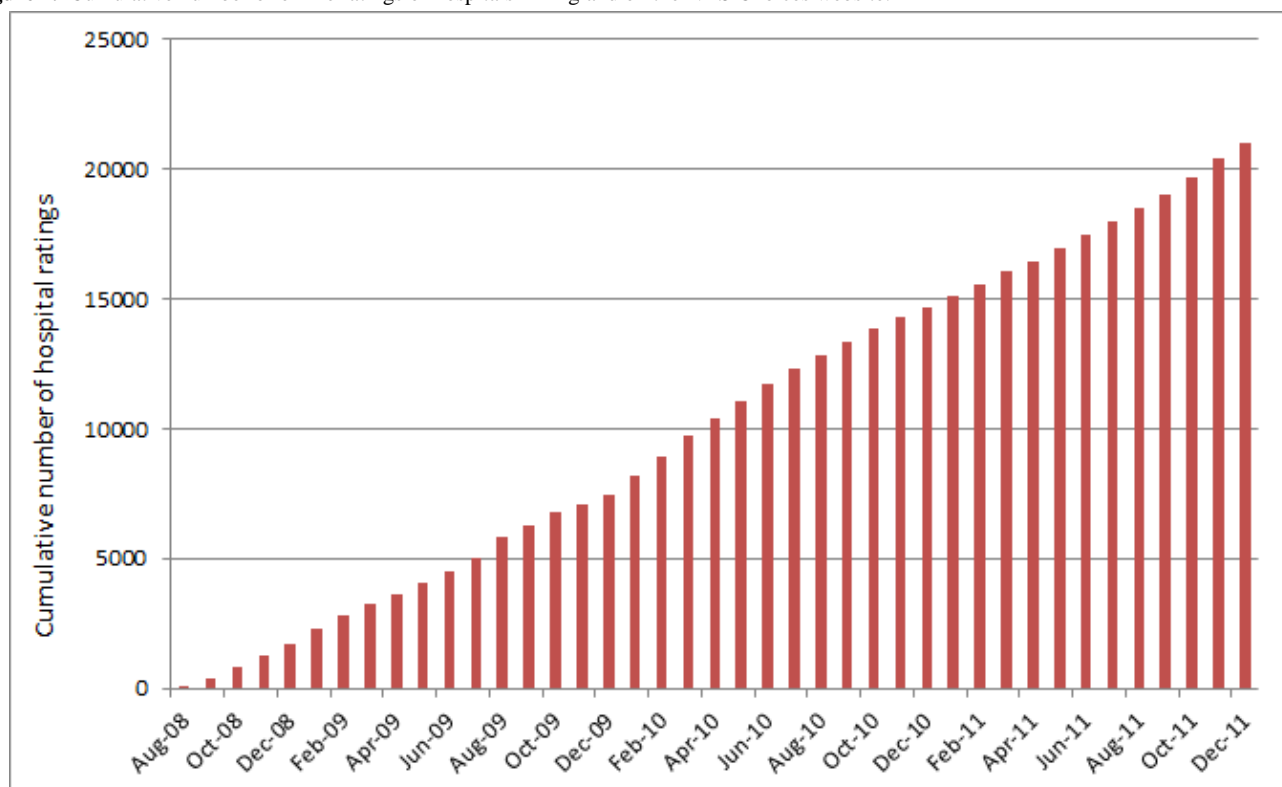
online reviews; quality transparency; public reporting

In a recent JMIR paper, Gao and colleagues demonstrated the growing number of internet-based ratings of physicians on a commercially-owned website in the USA [1].

In England, in keeping with our National Health Service, we have a government run website that allows patients to rate and comment on their care online in a similar way, but at the level of healthcare provider organisations rather than individuals. The website is called NHS Choices [2]. Gao suggests that their results demonstrate a positive correlation between online ratings and physician quality. We have similarly demonstrated how better online ratings at the organisational level are associated with better clinical outcomes [3], and patient experience measured by surveys [4] in England.

In a new analysis to allow comparison with Gao's results, we looked at the number of ratings of hospitals posted on the NHS Choices website over the period since it started (August 2008) to the end of 2011. There were 20,996 ratings of hospitals over the 40 month period, fewer than in the US. We found a more

gradual, linear increase in ratings in England (Figure 1) compared with the accelerating growth in ratings seen on commercial sites in the US [1]. We are not sure why the frequency of ratings is stable in England, but not increasing at the same rate as in the US. This may be because marketing budgets are lower for an English government run service compared to the more commercial advertising approach of US websites, leading to lower awareness of the websites in England. Alternatively, patients in England may be less culturally familiar with the concept of provider choice in healthcare, as the ability to choose between providers has only been introduced relatively recently in the English NHS while it may be a cultural norm in the US. This might result in English patients being less inclined to rate their care. We hope this adds to the work of our American colleagues, and demonstrates that the increasing number of online ratings of healthcare is an international phenomenon, even if England is perhaps at an earlier stage on the curve than the US.

Figure 1. Cumulative number of online ratings of hospitals in England on the NHS Choices website.

References

1. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res* 2012;14(1):e38 [FREE Full text] [doi: [10.2196/jmir.2003](https://doi.org/10.2196/jmir.2003)] [Medline: [22366336](https://pubmed.ncbi.nlm.nih.gov/22366336/)]
2. Department of Health. NHS Choices website. 2012. URL: <http://www.nhs.uk/Pages/HomePage.aspx> [accessed 2012-05-01] [WebCite Cache ID 67KZHKWkH]
3. Greaves F, Pape UJ, King D, Darzi A, Majeed A, Wachter RM, et al. Associations between Web-based patient ratings and objective measures of hospital quality. *Arch Intern Med* 2012 Mar 12;172(5):435-436. [doi: [10.1001/archinternmed.2011.1675](https://doi.org/10.1001/archinternmed.2011.1675)] [Medline: [22331980](https://pubmed.ncbi.nlm.nih.gov/22331980/)]
4. Greaves F, Pape UJ, King D, Darzi A, Majeed A, Wachter RM, et al. Associations between internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. *BMJ Qual Saf* 2012 Jul;21(7):600-605. [doi: [10.1136/bmjqs-2012-000906](https://doi.org/10.1136/bmjqs-2012-000906)] [Medline: [22523318](https://pubmed.ncbi.nlm.nih.gov/22523318/)]

Edited by G Eysenbach; submitted 04.05.12; peer-reviewed by G Gao, S Patel; comments to author 14.06.12; revised version received 14.06.12; accepted 27.06.12; published 29.06.12.

Please cite as:

Greaves F, Millett C
Consistently Increasing Numbers of Online Ratings of Healthcare in England
J Med Internet Res 2012;14(3):e94
URL: <http://www.jmir.org/2012/3/e94/>
doi:[10.2196/jmir.2157](https://doi.org/10.2196/jmir.2157)
PMID:[22742977](https://pubmed.ncbi.nlm.nih.gov/22742977/)

©Felix Greaves, Christopher Millett. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.06.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic

information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>