Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 14 (2012), Issue 1 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Review

Viewpoint

Letter

Corrigenda and Addendas

Original Paper

Scope of Policy Issues in eHealth: Results From a Structured Literature Review

Shariq Khoja[1,2], MBBS, MSc, PhD; Hammad Durrani[3], MBBS, MSc; Parvez Nayani[4], MBBS, MPH; Ammad Fahim[5], MBBS

[1]AKDN eHealth Resource Centre, The Aga Khan University, Nairobi, Kenya

[2]Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada

[3]VP Health & Operational Services Department, The Aga Khan University, Kabul, Afghanistan

[4]Programs in Afghanistan, The Aga Khan University, Kabul, Afghanistan

[5]Department of Community Health Sciences, The Aga Khan University, Karachi, Pakistan

Corresponding Author:
Shariq Khoja, MBBS, MSc, PhD
AKDN eHealth Resource Centre
The Aga Khan University
3rd Parklands Avenue
Nairobi, PO Box 30270-00100
Kenya
Phone: 254 20 3662236
Fax: 254 20 3744956
Email: shariq.khoja@aku.edu

Abstract

Background: eHealth is widely used as a tool for improving health care delivery and information. However, distinct policies and strategies are required for its proper implementation and integration at national and international levels.

Objective: To determine the scope of policy issues faced by individuals, institutions, or governments in implementing eHealth programs.

Methods: We conducted a structured review of both peer-reviewed and gray literature from 1998–2008. A Medline search for peer-reviewed articles found 40 papers focusing on different aspects of eHealth policy. In addition, a Google search found 20 national- and international-level policy papers and documents. We reviewed these articles to extract policy issues and solutions described at different levels of care.

Results: The literature search found 99 policy issues related to eHealth. We grouped these issues under the following themes: (1) networked care, (2) interjurisdictional practice, (3) diffusion of eHealth/digital divide, (4) eHealth integration with existing systems, (5) response to new initiatives, (6) goal-setting for eHealth policy, (7) evaluation and research, (8) investment, and (9) ethics in eHealth.

Conclusions: We provide a list of policy issues that should be understood and addressed by policy makers at global, jurisdictional, and institutional levels, to facilitate smooth and reliable planning of eHealth programs.

(J Med Internet Res 2012;14(1):e34) doi:10.2196/jmir.1633

KEYWORDS

eHealth; eHealth policies; telehealth; telemedicine; health informatics; electronic health records; health telematics; guidelines; standards

Introduction

eHealth policy can be defined as “a set of statements, directives, regulations, laws, and judicial interpretations that direct and manage the life cycle of eHealth” [1]. Recognition is growing in both developed and developing countries that eHealth is an important tool to reduce discrimination based on lack of access to information and to provide timely responses to matters affecting both personal and community health [2,3]. However, the use of eHealth within or between institutions involves several factors that require proper planning, supported by well-defined policies, rules, standards, or guidelines at the institutional,

jurisdictional, and global levels. The absence of these policies may cause problems during the cycle of eHealth planning that may lead to failures in achieving the intended goals. As a result, there could be inadvertent widening of gaps in health status and knowledge levels between different sectors of the population, and increasing rather than decreasing health inequity, also termed the digital divide [4,5]. Experience from the developed world shows that the most common barriers to successful eHealth planning include lack of information on the role of eHealth in the provision of health care, lack of operational and support policies, lack of demonstrated cost effectiveness, and lack of clinical proponents [6].

To complement the need for eHealth policies and strategies within countries, pressure is also developing at the global level for eHealth policies. The World Health Assembly (WHA) resolution of 2005, WHA 58.28, calls on member states to draw up long-term strategic plans for the development and implementation of eHealth. Thus, it is important for the planners of eHealth at different levels to develop policies that could facilitate the adoption of eHealth and prove its success through improvement in services and change in the health status of the population. It is important for global forces, governments, and institutional leadership to understand the range of policy issues that must be addressed at different levels and stages of an eHealth program to facilitate its planning and implementation.

The objective of this study was to conduct a detailed review of the literature to determine the scope of policy issues faced by individuals, institutions, or governments in implementing eHealth programs. The study does not recommend any policies or suggest the importance of any of the policy issues over the others.

Methods

We conducted a structured review of both peer-reviewed and gray literature. The search was conducted using the keywords *eHealth*, *telehealth*, *telemedicine*, *health informatics*, *electronic health records*, *health telematics*, *guidelines*, *policies*, *rules*, and *plans*. We collected the relevant information through the following process.

Review of Peer-Reviewed Literature

We searched PubMed using the above-mentioned key words. We chose only English-language articles published in peer-reviewed journals during the 10-years period 1998–2008. The search yielded 950 articles. After removing duplicates and articles beyond the scope of this study, we selected 150 articles for the review. We developed our own lists of policy issues, which we used for the abstracts and full-text review. The review was conducted in two stages: (1) two of the researchers reviewed 150 abstracts to select the articles that were relevant and merited a review of the full paper according to the checklist, and (2) the same two reviewers then reviewed 40 full papers that focused on different aspects of eHealth policy, or highlighted the policy issues in eHealth implementation.

Review of the Gray Literature:

We found 20 national- and international-level policy papers and documents through a Google search using the same key words described above. These articles were reviewed by two researchers to extract the policy issues and solutions described at different levels of care.

The list of issues was revised after the review. These issues were grouped into categories and themes for better understanding.

Results

We extracted 99 policy issues related to eHealth from the literature. These issues were grouped under 9 themes on the basis of similarities in their application. We identified the following themes for eHealth policies: (1) networked care, (2) interjurisdictional practice, (3) diffusion of eHealth/digital divide, (4) eHealth integration with existing systems, (5) response to new initiatives, (6) goal-setting for eHealth policy, (7) evaluation and research, (8) investment, and (9) ethics in eHealth.

eHealth policy issues were also divided on the basis of the levels where policies should be developed to deal with a particular issue. The levels identified for the policy development were global, jurisdictional (national or provincial/subnational), and individual institutions. We used the following operational definitions for these levels: (a) global: this level deals with the policies of global complementarity, such as standardization and interjurisdictional care, (b) jurisdictional (national and provincial/subnational): this level deals with the policies required to facilitate care within a health jurisdiction—that is, national or provincial/subnational governments, and (c) individual Institutions: this level deals with the policies required to facilitate eHealth at the local level—that is, individual institution or practice.

Below we describe the distribution of eHealth policy issues according to the themes and the levels of policy development.

Networked Care

The networked care theme [7-11] includes policy categories and issues that can enhance the ability of providers, departments, organizations, and jurisdictions to work in a coordinated environment to improve care of the population. Networked care covers the issues of interoperability [12,13], standardization [13], and intellectual property rights on material produced as a result of networked services [14], which need to be dealt with at the global level [12-14]. This theme also covers issues related to the use of acceptable, user-friendly, affordable, and reliable technology [14], to commitment to initial and ongoing funding [14,15], and to establishing local guidelines about sharing of information, standardization, communication, and control of malpractice [14-17], which can be dealt with at the jurisdictional level. The theme of networked care also includes issues related to change management, such as distribution of user workloads, improvement in readiness at the individual and institutional levels, and selection of simple and user-friendly technologies; financial matters, such as insurance requirements and reimbursement; guidelines related to sharing of information, knowledge, and services; cultural issues around communication and networking; and risk management [8,12-24]. Table 1

presents the matrix for the distribution of eHealth issues against the different levels under the theme of networked care.

Table 1. Networked care.

Level	Policy category	Issues	
a)	Global eHealth policies	i.	Functional and semantic interoperability
		ii.	Standardization of EHR[a]
		iii.	Intellectual property rights
b)	Jurisdictional (national and provincial/sub-national) policies	i.	Regulation of appropriate technologies
		ii.	Commitment of funds
		iii.	Standardization of EHR
		iv.	Sharing of services
		v.	Proper connectivity
		vi.	Control of malpractice
		vii.	Cultural issues in communication
c)	Institutional/individual policies	i.	Proper distribution of human resources
		ii.	Readiness building and effective change management
		iii.	Deployment of appropriate technologies
		iv.	Meeting the needs of insurance companies
		v.	Reimbursement and remuneration
		vi.	Sharing of patient information
		vii.	Sharing of knowledge
		viii.	Sharing of services
		ix.	Standardization measures for EHR
		x.	Ensuring integrity and quality of data and information
		xi.	Proper connectivity
		xii.	Risk management
		xiii.	Cultural issues in communication

[a] Electronic health record.

Interjurisdictional Practice

The interjurisdictional practice theme [8] includes policy categories and issues that deal with the transfer of information and provision of care between different jurisdictions. Interjurisdictional practice includes issues related to management of health information in shared environments [25,26], policies for privacy, confidentiality, and intellectual property rights [1,8,13,14,27], and guidelines for sharing knowledge and services [8,16], which can be dealt with globally. Interjurisdictional practice also deals with policies at the jurisdictional and individual levels, such as liability of care [8,12,14], proper licensing of health care providers [8,12,28], accreditation of individuals and institutions, and the defining of processes for coordinated services [1,8]. Table 2 shows the policy issues that encompasses the theme of interjurisdictional practice.

Table 2. Interjurisdictional practice.

Level	Policy category	Issues	
a)	Global eHealth policies	i.	Policies on managing health information on the Internet
		ii.	Intellectual property rights
		iii.	Complementarity of policies and health care regulations in different regions
		iv.	Sharing of knowledge
b)	Jurisdictional (national and provincial/sub-national) policies	i.	Accountability/liability of care
		ii.	Licensing
		iii.	Accreditation of services
		iv.	Local, national, and international policies
c)	Institutional/individual policies	i.	Accountability/liability of care

Diffusion of eHealth and Addressing the Digital Divide

The diffusion of eHealth [29,30] and digital divide [8,31,32] theme includes policy categories and issues that enhance the use of eHealth among populations who most need improved health services. These include policies and guidelines to allow greater penetration of telecommunication companies, such as mobile companies, Internet service providers, integrated services digital network providers, and satellite vendors, to reach the poorest countries [8,13], reduce the cost of telecommunication [27], provide universal and unlimited access to the Internet [33], and allow for appropriate use of eHealth for commercial and humanitarian purposes [34]. Other policy issues with a jurisdictional and individual focus, such as encouraging development and use of open-source technologies, increasing access to technology, reducing cost, and building local capacity [35], are also included under this theme. Table 3 shows the list of policy issues covered under the theme of diffusion of eHealth.

Table 3. Diffusion of eHealth/digital divide.

Level	Policy category	Issues	
a)	Global eHealth policies	i.	Telecommunication policies allowing increased access
		ii.	Control of technology costs
		iii.	Provision of universal and unlimited access to the Internet
		iv.	Humanitarian vs commercial policies
		v.	Sharing of knowledge and services
b)	Jurisdictional (national and provincial/subnational) policies	i.	Increasing focus on open-source technologies
		ii.	Telecommunication policies allowing increased access
		iii.	Control of technology costs
		iv.	Capacity building
c)	Institutional/individual policies	i.	Capacity building

Integration With Existing Systems

The theme of integration with existing systems [8,29] includes policy categories and issues that enable integration of eHealth projects and programs into regular services. The theme of integration includes jurisdictional policy issues such as setting targets for increasing interaction between different groups of providers and users, introducing decision-support systems to reduce the chance of errors [29], improving quality of care through eHealth and creating a learning environment, and changing business rules and models for integrating eHealth [33]. This theme also includes policies at the institutional level, such as defining the roles and responsibilities of different players [24], and creating guidelines on issues such as access to different gender and sociocultural groups, transfer and storage of information, patient consent, confidentiality, and privacy, which will help integrate eHealth with regular services [29,36].Table 4 lists the policy issues under the theme of integration with existing systems.

Table 4. eHealth integration with the existing systems.

Level	Policy category	Issues	
a)	Jurisdictional (national and provincial/subnational) policies	i.	Improvement of clinical effectiveness
		ii.	Improvement of quality of care
		iii.	Change in business rules in organizations
b)	Institutional/individual policies	i.	Redefinition of the roles and responsibilities of different players
		ii.	Wider ethical acceptability

Response to New Initiatives

The response to new initiatives [29] theme includes policy categories and issues that can enhance the capability of institutions to implement eHealth successfully. This theme includes jurisdictional policy issues, such as guidelines for identifying and including stakeholders from different user and support groups in the planning of eHealth programs [24]. This theme also covers policy issues at the institutional level, such as defining the roles and responsibilities of different players such as local providers and specialists [24], defining the processes for change management [29], ensuring training and support to all users [29], defining the rules for procurement of equipment [13], distribution of bandwidth, and distribution and security of wireless networks [37], maintaining doctor–patient relationship [38], and evaluating new technologies in local environments before implementation to avoid difficulties and failure [29]. Table 5 lists the policy issues under the theme of response to new initiatives.

Table 5. Response to new initiatives.

Level	Policy category	Issues	
a)	Jurisdictional (national and provincial/subnational) policies	i.	Definition of stakeholders at different levels
b)	Institutional/individual policies	i.	Definition of the roles and responsibilities of different players, such as local providers and specialists
		ii.	Change management
		iii.	End-user support
		iv.	Regulation of information technology use
		v.	Maintenance of the doctor–patient relationships
		vi.	Wireless networks and security issues
		vii.	Evaluation of new technologies in local environments

Policy Goal-Setting

The policy goal-setting [39] theme includes policy categories and issues that can guide the process of defining policies for eHealth. Key global considerations in this regard include recognition of eHealth as part of the broader development effort, in terms of assisting national health systems and recognizing eHealth as part of the global health agenda [40], and encouraging a global commitment for funding eHealth programs. This theme also includes jurisdictional considerations, such as developing policies to encourage growth of the telecommunications sector and to increase connectivity in remote areas [41]; increasing flexibility between governments and private institutions to align with changing information technology environments and policies [13]; encouraging innovation and development [38]; covering the costs of equipment and time needed for health care providers to bring eHealth services into broad acceptance [40]; and developing governance and management structures [41,42]. Institutional policy issues, such as ensuring universal standards of care, and allotting and distributing the workload for health care providers and technical and managerial staff [43], are also included under this theme. Table 6 lists the policy issues under the theme of goal-setting for eHealth policy.

Table 6. Goal-setting for eHealth policy.

Level	Policy Categories	Issues	
a)	Global eHealth policies	i.	Integration of eHealth into the overall development effort
		ii.	Funding of eHealth programs
b)	Jurisdictional (national and provincial/subnational) policies	i.	Provision of suitable telecommunications infrastructure to promote eHealth
		ii.	Alignment of policies with information technology innovations
		iii.	Innovative and forward-looking policies
		iv.	Coverage of the opportunity cost of health providers' time
		v.	Timing of government action
		vi.	Development of leadership structures for eHealth programs
		vii.	Development of strategies for eHealth adoption
		viii.	Information governance
c)	Institutional/individual policies	i.	Standards of care
		ii.	Guidelines for human resources

Evaluation and Research

The evaluation and research [39] theme includes policy categories and issues that can guide the process of evaluation and research to generate evidence for the adoption of eHealth. These policy issues include measurement of the time spent during teleconsultations and justification of the resources spent on setting up eHealth services [44], cost effectiveness [45], impact on health care management [45], demonstration of improvement in health outcomes [46], and enhancement of clinical effectiveness [38] and learning [3]. Other issues at the level of individual institutions include providing an environment for testing and simulating eHealth initiatives [37], encouraging interdisciplinary research [47], and disseminating results for policy making and the benefit of users [3]. Table 7 lists the policy issues under the theme of evaluation and research.

Table 7. Evaluation and research.

Level	Policy category	Issues	
a)	Jurisdictional (national and provincial/subnational) policies	i.	Justification of health providers' time
		ii.	Cost effectiveness
		iii.	Impact of eHealth on health care management
		iv.	Demonstration of health outcomes
		v.	Evidence of clinical effectiveness
		vi.	Progress in learning
b)	Institutional/individual policies	i.	Provision of simulation environment
		ii.	Encouragement of coordinated research
		iii.	Dissemination for policy making and benefit of others

Investment

The investment [9] theme includes policy issues that can suggest business models for eHealth adoption. This theme includes encouraging the use of eHealth by health care institutions to increase the number of clients and to grow their businesses [14] and encouraging partnerships between public and private institutions, or within the same sector. It also includes cross-jurisdictional advertisement and sale of drugs and services. Table 8 lists the policy issues under the theme of investment.

Table 8. Investment.

Level	Policy category	Issues	
a)	Jurisdictional (national and provincial/subnational) policies	i.	Use of eHealth for commercial purposes
		ii.	Public–private partnership
		iii.	Cross-border advertisement and sale of drugs

Ethics and Legal Issues in eHealth

The theme of ethics and legal issues in eHealth [25] includes the ethical issues that must be considered during adoption of eHealth. These include global policy issues, such as managing health information on the Internet [25,26] and ensuring privacy of health information [27]. This theme also includes jurisdictional and institutional policy issues, such as patient consent [24,48], liability of care [45,48-50], medicolegal issues [45], patients' rights to access their own health information [14], security of information during portability [48], maintenance of quality of care [14,17], and cultural issues in communication [14]. Table 9 lists the policy issues under the theme of ethics and legal issues in eHealth.

Table 9. Ethical and legal issues in eHealth.

Level	Policy category	Issues	
a)	Global eHealth policies	i.	Management of health information on the Internet
		ii.	Health information privacy
b)	Jurisdictional (national and provincial/subnational) policies	i.	Consent for care in eHealth
		ii.	Liability issues (medical malpractice liability)
		iii.	Medicolegal issues
		iv.	Patients' right to access information
		v.	Security of information during portability
		vi.	Control of malpractice
		vii.	Cultural issues in communication

Discussion

This policy paper provides a spectrum of eHealth issues that require policies for different levels of decision makers. It is important for the policy makers at the global, national, and institutional levels to understand the scope and importance of these issues; to analyze their current situation; and to take a proactive approach to developing policies that facilitate smooth and reliable planning of eHealth programs.

Based on our findings, we recommend a combination of policies at different levels when developing eHealth policies. Many strategies suggest that the development of supportive policies should be part of the eHealth strategies of countries and organizations. These recommendations should, however, come from the user groups and managers of eHealth programs in each country. It is therefore important to increase the awareness that health care providers, managers, and policy makers at all levels have of eHealth policy issues by providing them guidelines and support to develop these policies.

The main strength of this paper is that it draws on policy issues from both peer-reviewed and gray literature. Our review of policy papers provided a detailed analysis of different policy issues and their relationship with other such issues at different levels. Due to the limitation of time and resources, it was not possible to review the literature in languages other than English, or to conduct a detailed systematic review.

As a follow-up to this study, an environmental scan at all three levels (global, national, and institutional) should be conducted to identify and study the already-existing policies on the issues identified in this paper. There is a need to study the successes and failures of these policies, which will support the development of guidelines for policy makers at the global, regional, national, and local levels. This would lead to policy formulation that not only benefits their own eHealth programs but also generates knowledge to support programs in other areas.

Finally, we have also developed a 2-way matrix that will show which policies are relevant at different stages of eHealth planning, implementation, integration, and sustainability. The findings of this exercise merit a separate publication, which should help eHealth planners to assess the need for relevant policies for different eHealth initiatives and to develop the policy issues most relevant to that stage.

Acknowledgments

We thank the World Health Organization and the Rockefeller Foundation for funding and supporting preparation of the manuscript. The policy paper was prepared for the Bellagio eHealth Workshop series. We also appreciate Syeda Suha Tirmizi for her help in drafting the manuscript.

Conflicts of Interest

None declared.

References

1. Scott RE, Chowdhury MF, Varghese S. Telehealth policy: looking for global complementarity. J Telemed Telecare 2002;8 Suppl 3:S3:55-S3:57. [Medline: 12661624]
2. World Health Organization, Regional Office for the Eastern Mediterranean. 2005. E-Health in the Eastern Mediterranean URL: http://www.emro.who.int/his/ehealth/AboutEhealth.htm [accessed 2010-06-23] [WebCite Cache ID 5qhCuy6o8]
3. WHO Secretariat. World Health Organization Executive Board. 2004 Dec 16. eHealth Report by the Secretariat URL: http://apps.who.int/gb/ebwha/pdf_files/EB115/B115_39-en.pdf [accessed 2010-06-23] [WebCite Cache ID 5qhDYET42]
4. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: eHealth literacy, health disparities, and the digital divide. Health Mark Q 2008;25(1-2):175-203. [doi: 10.1080/07359680802126301] [Medline: 18935884]
5. Bashshur R, Shannon G, Sapci AH. The eHealth digital divide: concept, implications, and solutions. In: Proceedings. 2003 Presented at: eHealth 2003: Third Annual Conference of International eHealth Association; October 16-17, 2003; London, UK URL: http://www.ehealthinternational.org/pdfs/Bashshur.pdf
6. Lam DM, Mackenzie C. Human and organizational factors affecting telemedicine utilization within U.S. military forces in Europe. Telemed J E Health 2005 Feb;11(1):70-78. [doi: 10.1089/tmj.2005.11.70] [Medline: 15785223]
7. Broens TH, Huis in't Veld RM, Vollenbroek-Hutten MM, Hermens HJ, van Halteren AT, Nieuwenhuis LJ. Determinants of successful telemedicine implementations: a literature study. J Telemed Telecare 2007;13(6):303-309. [doi: 10.1258/135763307781644951] [Medline: 17785027]
8. Scott RE, Lee A. E-health and the Universitas 21 organization: 3. Global policy. J Telemed Telecare 2005;11(5):225-229. [doi: 10.1258/1357633054471867] [Medline: 16035963]
9. Chandrasekhar CP, Ghosh J. Information and communication technologies and health in low income countries: the potential and the constraints. Bull World Health Organ 2001;79(9):850-855 [FREE Full text] [Medline: 11584733]
10. eHealth ERA Team. European Commission, Information Society and Media Directorate General. 2007 Mar. eHealth Priorities and Strategies in European Countries: eHealth ERA report URL: http://www.ehealtheurope.net/img/document_library0282/ehealthPriorities_and_Strategies.pdf [accessed 2012-01-22] [WebCite Cache ID 5qfcIJQv4]
11. Donahue M. Regulating telehealth in Ontario--next step in the transformation agenda. Health Law Can 2006 Aug;27(1):5-28. [Medline: 17444377]
12. Shannon G, Nesbitt T, Bakalar R, Kratochwill E, Kvedar J, Vargas L. Telemedicine/telehealth: an international perspective. Organizational models of telemedicine and regional telemedicine networks. Telemed J E Health 2002;8(1):61-70. [doi: 10.1089/15305620252933400] [Medline: 12020406]
13. Cunningham R. Old before its time: HIPAA and e-health policy. Health Aff (Millwood) 2000 Dec;19(6):231-238 [FREE Full text] [Medline: 11192408]
14. Scott RE, Jennett P, Yeo M. Access and authorisation in a Glocal e-Health Policy context. Int J Med Inform 2004 Mar 31;73(3):259-266. [doi: 10.1016/j.ijmedinf.2003.11.020] [Medline: 15066556]
15. Khoja S, Scott RE, Casebeer AL, Mohsin M, Ishaq AF, Gilani S. e-Health readiness assessment tools for healthcare institutions in developing countries. Telemed J E Health 2007 Aug;13(4):425-431. [doi: 10.1089/tmj.2006.0064] [Medline: 17848110]
16. Schmeida M, McNeal R, Mossberger K. Policy determinants affect telehealth implementation. Telemed J E Health 2007 Apr;13(2):100-107. [doi: 10.1089/tmj.2006.0017] [Medline: 17489696]
17. Silverman RD. Current legal and ethical concerns in telemedicine and e-medicine. J Telemed Telecare 2003;9 Suppl 1:S67-S69. [doi: 10.1258/135763303322196402] [Medline: 12952729]
18. Hudson HE. Rural telemedicine: lessons from Alaska for developing regions. Telemed J E Health 2005 Aug;11(4):460-467. [doi: 10.1089/tmj.2005.11.460] [Medline: 16149892]
19. Khoja S, Scott RE, Ishaq F, Mohsin M. Reliability testing of eHealth readiness assessment tools. eHealth Int J 2007;3(1):31-37.
20. Khoja S, Casebeer A, Scott RE, Gilani SN. Validating eHealth readiness assessment tools by using qualitative research methods. eHealth Int J 2007;3(1):24-30.
21. Gray GA, Stamm BH, Toevs S, Reischl U, Yarrington D. Study of participating and nonparticipating states' telemedicine Medicaid reimbursement status: Its impact on Idaho's policymaking process. Telemed J E Health 2006 Dec;12(6):681-690. [doi: 10.1089/tmj.2006.12.681] [Medline: 17250490]
22. Grigsby B, Brega AG, Bennett RE, Devore PA, Paulich MJ, Talkington SG, et al. The slow pace of interactive video telemedicine adoption: the perspective of telemedicine program administrators on physician participation. Telemed J E Health 2007 Dec;13(6):645-656. [doi: 10.1089/tmj.2007.0090] [Medline: 18092926]
23. Doarn CR, Merrell RC. Telemedicine and e-Health for international medical issues. Telemed J E Health 2005 Dec;11(6):621-623. [doi: 10.1089/tmj.2005.11.621] [Medline: 16430379]
24. Nerlich M, Balas EA, Schall T, Stieglitz SP, Filzmaier R, Asbach P, et al. Teleconsultation practice guidelines: report from G8 Global Health Applications Subproject 4. Telemed J E Health 2002;8(4):411-418. [doi: 10.1089/15305620260507549] [Medline: 12626110]
25. Eysenbach G. Towards ethical guidelines for e-health: JMIR theme issue on eHealth ethics. J Med Internet Res 2000;2(1):E7 [FREE Full text] [doi: 10.2196/jmir.2.1.e7] [Medline: 11720926]

26. Powell JA, Lowe P, Griffiths FE, Thorogood M. A critical analysis of the literature on the Internet and consumer health information. J Telemed Telecare 2005;11 Suppl 1:41-43. [doi: 10.1258/1357633054461642] [Medline: 16035990]
27. Jacobson PD, Selvin E. Licensing telemedicine: the need for a national system. Telemed J E Health 2000;6(4):429-439. [doi: 10.1089/15305620050503915] [Medline: 11242552]
28. Goldberg MA, Sharman Z, Bell B, Ho K, Patil N. E-health and the Universitas 21 organization: 4. Professional portability. J Telemed Telecare 2005;11(5):230-233. [doi: 10.1258/1357633054471894] [Medline: 16035964]
29. Bower DJ, Barry N, Reid M, Norrie J. Designing and implementing E-health Applications in the UK's National Health Service. J Health Commun 2005 Dec;10(8):733-750. [doi: 10.1080/10810730500326732] [Medline: 16316936]
30. See A, Lim AC, Le K, See JA, Shumack SP. Operational teledermatology in Broken Hill, rural Australia. Australas J Dermatol 2005 Aug;46(3):144-149. [doi: 10.1111/j.1440-0960.2005.00166.x] [Medline: 16008643]
31. Doarn CR, Merrell RC. A national strategy for telemedicine and e-health. Telemed J E Health 2007 Jun;13(3):243-244. [doi: 10.1089/tmj.2007.9982] [Medline: 17603825]
32. Marshall S, Taylor W, Yu Z. Closing the Digital Divide: Transforming Regional Economies and Communities With Information Technology. Westport, CT: Praeger; 2003.
33. Lim AC, Egerton IB, Shumack SP. Australian teledermatology: the patient, the doctor and their government. Australas J Dermatol 2000 Feb;41(1):8-13. [Medline: 10715894]
34. Varghese S, Scott RE. Categorizing the telehealth policy response of countries and their implications for complementarity of telehealth policy. Telemed J E Health 2004;10(1):61-69. [doi: 10.1089/153056204773644599] [Medline: 15104917]
35. Gagnon MP, Duplantie J, Fortin JP, Landry R. Exploring the effects of telehealth on medical human resources supply: a qualitative case study in remote regions. BMC Health Serv Res 2007;7:6 [FREE Full text] [doi: 10.1186/1472-6963-7-6] [Medline: 17217534]
36. Stroetmann VN, Hüsing T, Kubitschke L, Stroetmann KA. The attitudes, expectations and needs of elderly people in relation to e-health applications: results from a European survey. J Telemed Telecare 2002;8 Suppl 2:82-84. [doi: 10.1258/135763302320302154] [Medline: 12217149]
37. Boyle J. Wireless technologies and patient safety in hospitals. Telemed J E Health 2006 Jun;12(3):373-382. [doi: 10.1089/tmj.2006.12.373] [Medline: 16796507]
38. Yellowlees P. Government relations, government regulations: jumping through the hoops. J Telemed Telecare 2002;8 Suppl 3:S3:83-S3:85. [Medline: 12661634]
39. Jennett PA, Scott RE, Affleck Hall L, Hailey D, Ohinmaa A, Anderson C, et al. Policy implications associated with the socioeconomic and health system impact of telehealth: a case study from Canada. Telemed J E Health 2004;10(1):77-83. [doi: 10.1089/153056204773644616] [Medline: 15104919]
40. Grigsby J, Rigby M, Hiemstra A, House M, Olsson S, Whitten P. Telemedicine/telehealth: an international perspective. The diffusion of telemedicine. Telemed J E Health 2002;8(1):79-94. [doi: 10.1089/15305620252933428] [Medline: 12020408]
41. Garfield MJ, Watson RT. Four case studies in state-supported telemedicine initiatives. Telemed J E Health 2003;9(2):197-205. [doi: 10.1089/153056203766437534] [Medline: 12855042]
42. Huston JL. Information governance standards for managing e-health information. J Telemed Telecare 2005;11 Suppl 2:S56-S58. [doi: 10.1258/135763305775124614] [Medline: 16375798]
43. MacDonald-Rencz S, Cradduck T, Parker-Taillon D. The national initiative for telehealth guidelines. Telemed J E Health 2004;10(1):113-114. [doi: 10.1089/153056204773644661] [Medline: 15104924]
44. McIntosh WA, Alston LT, Booher JR, McCaughan W, Segura CB, Sykes D, et al. Time spent with patients and charges to patients for specialty consultations using telemedicine. Telemed J E Health 2003;9(4):345-350. [doi: 10.1089/153056203772744662] [Medline: 14980091]
45. Lacroix A, Lareng L, Padeken D, Nerlich M, Bracale M, Ogushi Y, et al. International concerted action on collaboration in telemedicine: recommendations of the G-8 Global Healthcare Applications Subproject-4. Telemed J E Health 2002;8(2):149-157. [doi: 10.1089/15305620260008084] [Medline: 12079604]
46. Wootton R. Telemedicine support for the developing world. J Telemed Telecare 2008;14(3):109-114. [doi: 10.1258/jtt.2008.003001] [Medline: 18430271]
47. Stroetmann KA, Stroetmann VN. Towards an interoperability framework for a European e-health research area: locating the semantic interoperability domain. 2005 Presented at: WHO/EC Workshop on Semantic Interoperability; Feb 14-15, 2005; Brussels, Belgium URL: http://www.who.int/classifications/terminology/stroetmann.pdf
48. Poropatich RK, DeTreville R, Lappan C, Barrigan CR. The U.S. Army telemedicine program: general overview and current status in Southwest Asia. Telemed J E Health 2006 Aug;12(4):396-408. [doi: 10.1089/tmj.2006.12.396] [Medline: 16942411]
49. Jennett P, Jackson A, Ho K, Healy T, Kazanjian A, Woollard R, et al. The essence of telehealth readiness in rural communities: an organizational perspective. Telemed J E Health 2005 Apr;11(2):137-145. [doi: 10.1089/tmj.2005.11.137] [Medline: 15857254]
50. Légaré E, Vincent C, Lehoux P, Anderson D, Kairy D, Gagnon MP, et al. Telehealth readiness assessment tools. J Telemed Telecare 2010;16(3):107-109. [doi: 10.1258/jtt.2009.009004] [Medline: 20386031]

Abbreviations

WHA: World Health Assembly

Edited by G Eysenbach; submitted 27.08.10; peer-reviewed by M Mars, R Scott, S Lim Choi Keung, C Hullin Lucay Cossio; comments to author 20.11.10; revised version received 22.07.11; accepted 25.08.11; published 17.02.12.

Please cite as:
Khoja S, Durrani H, Nayani P, Fahim A
Scope of Policy Issues in eHealth: Results From a Structured Literature Review
J Med Internet Res 2012;14(1):e34
URL: http://www.jmir.org/2012/1/e34/
doi:10.2196/jmir.1633
PMID:22343270

Original Paper

Identifying Factors for Optimal Development of Health-Related Websites: A Delphi Study Among Experts and Potential Future Users

Francine Schneider[1], MSc; Liesbeth van Osch[1], PhD; Hein de Vries[1], PhD

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

Corresponding Author:
Francine Schneider, MSc
CAPHRI
Department of Health Promotion
Maastricht University
P.O. Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433882397
Fax: 31 33671032
Email: francine.schneider@maastrichtuniversity.nl

Abstract

Background: The Internet has become a popular medium for offering tailored and targeted health promotion programs to the general public. However, suboptimal levels of program use in the target population limit the public health impact of these programs. Optimizing program development is considered as one of the main processes to increase usage rates.

Objective: To distinguish factors potentially related to optimal development of health-related websites by involving both experts and potential users. By considering and incorporating the opinions of experts and potential users in the development process, involvement in the program is expected to increase, consequently resulting in increased appreciation, lower levels of attrition, and higher levels of sustained use.

Methods: We conducted a systematic three-round Delphi study through the Internet. Both national and international experts (from the fields of health promotion, health psychology, e-communication, and technical Web design) and potential users were invited via email to participate. During this study an extensive list of factors potentially related to optimal development of health-related websites was identified, by focusing on factors related to layout, general and risk information provision, questionnaire use, additional services, and ease of use. Furthermore, we assessed the extent to which experts and potential users agreed on the importance of these factors. Differences as well as similarities among experts and potentials users were deduced.

Results: In total, 20 of 62 contacted experts participated in the first round (32% response rate); 60 of 200 contacted experts (30% response rate) and 210 potential users (95% response rate) completed the second-round questionnaire, and 32 of 60 contacted experts completed the third round (53% response rate). Results revealed important factors consented upon by experts and potential users (eg, ease of use, clear structure, and detailed health information provision), as well as differences regarding important factors consented upon by experts (eg, visual aids, self-monitoring tool, and iterative health feedback) or by potential users only (eg, bread crumb navigation and prevention of receiving spam).

Conclusions: This study is an important first step in determining the agreed-upon factors that should be taken into account when developing online health promotion programs. The public health impact of these programs will be improved by optimizing the development process in line with these factors.

(J Med Internet Res 2012;14(1):e18) doi:10.2196/jmir.1863

KEYWORDS

Health promotion; Internet interventions; health behavior change; optimal development; Delphi study; experts; users

Introduction

Worldwide, more and more people are accessing the Internet in search of health-related information [1]. It is estimated that globally a minimum of nearly seven million health-related Internet searches are conducted daily [2]. Since Internet penetration rates are still expanding, with currently almost two billion people having access to the Internet, the number of health-related searches is also expected to increase [3]. Therefore, the Internet is a promising channel for offering a broad range of health-related information, such as background information on health, treatment information, medication information, and health behavior information [1].

Due to this high level of accessibility and its potential to reach large numbers of people [3], the Internet has also become a popular medium in the field of health promotion for offering tailored and targeted health promotion programs [4,5]. As a consequence, in recent years positive effects of online interventions applying computer-tailored techniques have been reported addressing different health behaviors [6], such as physical activity [7,8], nutrition [9,10], smoking cessation [11-14], and alcohol consumption [15,16]. Although these tailored interventions are very promising and have proven to be effective, actual reach is failing to live up to the high expectations [17-22]. Since the public health impact of interventions is determined not only by their efficacy but also by their levels of exposure in the target group [23,24], it is imperative to put effort into optimizing the level of exposure to Internet-delivered lifestyle interventions.

Successful exposure is partly defined by the level of first-time access of the intervention, also referred to as *first use* or adoption ([19,24,25]. Besides first use, *prolonged use* of the intervention is essential. That is, engaging users in the intervention for a substantial amount of time fosters their knowledge of its content and involvement in its effective components, which consequently increases the chances of health behavior change [17,26]. Since health behavior change is a complex and continuous process, achieving sustained behavior change depends on both the intensity of the intervention and the number of times the intervention is visited [27,28]. Due to this high dose–response relationship, ensuring adherence or *sustained use* of the program is essential to further maximize its effect on subsequent health behavior change [29,30]. Both prolonged and sustained use of the intervention can be influenced not only by user characteristics (eg, demographic characteristics and motivation to use the intervention [17,31,32]) but also by specific strategies to increase adherence (eg, sending periodic reminders [33,34]) and by intervention characteristics (eg, appearance and content of the intervention [33,35]).

As online behavior change interventions are delivered to the public by using a website or Web-based program, development of the website or program refers to composition of the actual intervention, and requires careful composition of the complete website or program it is embedded in. The website as a mode of delivery is described in the Internet intervention model as consisting of 8 main areas [35]: appearance (eg, the organization of information), behavioral prescriptions (eg, instructions on how to achieve behavior change), burdens (eg, poor navigation applications), content (eg, treatment information), delivery (eg, use of animations, audio, or graphics), message (eg, credibility and likability of the source), participation (eg, degree of interaction or the use of rewards), and assessment (eg, measuring needs of users or adjusting content to personal wishes). Adjusting these characteristics enables tailoring of the website to special needs of the population under consideration. All individual characteristics in these areas should therefore be kept in mind while developing the program.

The primary characteristics of an intervention are determined during its development process. According to diffusion theory and social marketing principles [25,36], program development is one of the main processes to influence adoption rates of a new product (eg, a new website or online intervention). Besides including experts in the development process, it should also be done in accordance with the needs and wishes of the target group [36]. By considering and incorporating the opinions of potential future users, involvement is expected to increase, consequently resulting in increased appreciation of the intervention, lower levels of attrition, and higher levels of prolonged and sustained use [37]. A new intervention should, therefore, be developed in close collaboration with the target group.

Although many studies have investigated the effectiveness and appreciation of numerous components of websites that deliver health behavior interventions, results have been inconclusive. Some studies recommend using a multimedia approach [38] or the use of interactive tools, such as surveys, quizzes, and games, whereas other studies seem to contradict these findings [5]. Systematically studying different elements of websites is a very extensive and time-consuming process. Even though some elements of websites should be incorporated at all times, such as accurate and comprehensible information [39], a lot of effort should still be put into making other elements operational [35].

To date, studies on development of health-related websites have not included perspectives of both experts and users. Hence, this study was a first attempt to identify which elements on health-related websites are perceived as necessary and preferable by these two groups. The current study thus included both experts in the fields of health promotion, health psychology, e-communication, and technical Web design and potential future users. By including elements that are perceived as important by both groups while developing the website, developers may optimize exposure rates of the program [36]. Furthermore, involving different groups of experts will lead to more diverse information regarding health communication and behavior change as well as technical information regarding website development. To identify the potential factors that are related to optimal development of health-related websites, we conducted a three-round Delphi study. The specific aim of the study was to identify factors that are associated with optimal development of health-related websites. Besides identifying these factors, we investigated the degree to which experts and potential future users agreed on the importance of the factors.

Methods

We conducted a three-round Delphi study through the Internet. Due to its systematic nature, a Delphi study is considered to be an appropriate method to derive consensus on health-related issues for which scientific evidence is incomplete or scarce by involving a representative panel of experts [40]. Also, because after each round feedback on group results is provided, the iterative approach allows participants to adjust their opinions when needed. Finally, the structure of the study guarantees anonymity of the participants, thereby preventing conformity biases [41,42].

The current Delphi study consisted of two substudies. The first substudy (study 1a), including only experts, consisted of three rounds. The first round aimed at providing a list of potential factors related to optimal development of health-related websites by means of an open-ended questionnaire. Next, for the second round, experts were invited to rate the importance of all factors identified in the first round by using a structured questionnaire. Finally, a third round enabled experts to reevaluate their opinions by providing controlled feedback regarding group mean scores, thereby producing consensus. The second substudy (study 1b) included only potential future users and consisted of one round. This round resembled the second round of the first substudy and allowed users to rate the importance of those factors that were identified by experts in the first round by using the same structured questionnaire. We compared second-round results from both studies in order to identify potential differences between experts and potential users regarding the importance they placed on factors related to optimal design.

Study 1a: Experts

First Round

Procedure and Participants

For the first round, we selected experts from the fields of health promotion, health psychology, e-communication, and technical Web design to obtain a variety of insights from researchers with both theory-based and more practice-based backgrounds. Invited experts with a theory-based background were all first or second authors on scientific papers in the field of eHealth and eHealth promotion published between 2000 and 2009. We used database searches in PsychINFO and Medline to identify experts and to examine reference lists from related papers, book chapters, review studies, and conference abstracts. We selected experts with a practice-based background on the basis of their publications, but also by approaching our own network and by asking responding experts to provide names of important experts in the field.

This resulted in a list of 62 experts who were invited by email to participate in all three rounds of the Delphi study. The email contained detailed information on the goal and study procedure, as well as a link referring them directly to the first-round questionnaire. Nonresponders received a reminder email after the 3-week response period expired. A total of 20 experts (32% response rate) responded to the invitation.

Questionnaire

The first-round questionnaire consisted of 7 open-ended questions. The questions pertained to different subjects related to the development of health-related websites: (1) optimal layout, (2) type of general information provided, (3) type of health risk information provided, (4) ease of use, (5) use of visual aids, and (6) additional information provided. Health-related websites often provide questionnaires to allow visitors the opportunity to assess their own current health status. To provide an accurate update of their health status these questionnaires often tend to be extensive and therefore sensitive to early dropout. Therefore, we also included one open-ended question to gain more insight into factors that contribute to completion of the questionnaires often provided on health-related websites (7). Health-related websites were defined as websites aiming at assisting people to adopt a healthier lifestyle, by offering them important and diverse information regarding health and health-related behaviors (eg, physical activity, smoking, alcohol consumption, and nutrition).

Data Analysis

Responses of experts were analyzed, resulting in an extensive list of potential factors related to optimal development of health-related websites. Two researchers independently listed all unique factors and combined similar responses into 1 factor. For those factors on which no agreement was obtained, a third researcher was approached to give a decisive answer.

Second Round

Procedure and Participants

Experts participating in the first round were also invited to participate in the second round. We selected an additional 180 experts by means of the same strategies used for the first round.

A total of 200 experts received an email inviting them to participate in the second and third rounds. The invitation contained a link that directed experts to the second-round questionnaire. Nonresponders received two reminder emails: after 3 and after 5 weeks. A total of 60 experts (30% response rate) responded to the invitation (Table 1).

Table 1. First-, second-, and third-round rates of responses of experts

Field	First round (n = 62)		Second round (n = 200)		Third round (n = 60)	
	Invitations	Response (%)	Invitations	Response (%)	Invitations	Response (%)
Health promotion	16	7 (44%)	60	18 (30%)	18	10 (56%)
Health psychology	16	5 (31%)	60	14 (23%)	14	8 (57%)
E-communication	15	5 (33%)	40	12 (30%)	12	5 (42%)
Technical Web design	15	2 (13%)	40	16 (40%)	16	9 (56%)
Total	62	20 (32%)	200	60 (30%)	60	32 (53%)

Questionnaire

The second-round questionnaire was composed based on the factors identified by experts in the first round. This resulted in a questionnaire consisting of 85 structured questions. Experts were asked to rate these factors on a 7-point Likert scale ranging from 1 (not at all important) to 7 (extremely important).

Data Analysis

The closed-ended questions were analyzed following the standards for analyzing data for a Delphi study, by calculating median scores, also referred to as the 50th percentile score, to determine the importance of the various factors. Furthermore, the interquartile range (IQR) was calculated to assess the degree of agreement between the experts on the importance of the factors [31,41,43]. The IQR represents the distance between the 25th and the 75th percentiles, with a small value indicating a higher degree of agreement. An IQR smaller than 1 is considered to indicate good consensus on a 7-point scale as used in the present study and means that more than 50% of all cases fall within 1 point of one another [44]. To deduce those factors that were considered to be either very or extremely important by the majority of experts (IQR < 1), the cut-off point for importance was a median score of ≥6.

Third Round

Procedures and Participants

All experts participating in the second round (n = 60) were invited to participate in the third and final Delphi round, using the same procedure as in the previous round. A total of 32 experts (53% response rate of second-round participants) responded to the invitation (Table 1).

Questionnaire

The third-round questionnaire was an adapted version of the second-round questionnaire, containing only those factors (n = 27) on which no consensus was obtained during the second round (IQR > 1). In line with the Delphi method, additional feedback on second-round group results (median and IQR) was provided, enabling experts to rerate their answers.

Data Analysis

The degree of agreement and consensus among experts was measured by computing median scores and IQRs.

Study 1b: Potential Future Users

Round 2

Procedures and Participants

Potential future users were recruited through an independent, commercial Dutch Internet research panel [45]. From this panel, consisting of approximately 20,000 members, a sample of 220 members were invited to participate in this study. Potential users were invited to take part only in the second round of this Delphi study, since the main goal was to compare users' opinions with experts' opinions and not to force consensus among users on the different factors. Respondents from this panel were rewarded for their participation in this study in accordance with the standards of the consumer panel (approximately €5). All participants received an email invitation informing them about the goal and content of the study. The invitation also contained a link that directed participants to the questionnaire. A total of 210 potential future users (95.5% response rate) responded to the invitation (Table 2).

Table 2. Demographic profile of potential future users (n = 210)

Demographic characteristic	%	n
Age (years)		
Range	19–65	
Mean (SD)	46.49 (47.00)	
Gender		
Male	51	106
Female	49	104
Ethnicity		
Native	98	206
Nonnative	2	4
Education level		
Low	11	24
Medium	59	123
high	30	63
Income		
Less than average	29	61
Average	26	54
More than average	45	94
Internet use: work or study [a]		
Frequent	55	116
Average	12	26
Infrequent	10	20
Not applicable	23	48
Internet use: private use [a]		
Frequent	96	202
Average	3	6
Infrequent	1	2
Internet use: health purposes		
Frequent	21	45
Average	69	144
Infrequent	10	21

[a] Frequent user: >3 times/week; average user: every week, but not >3 days/week; infrequent user: <1/week.

Questionnaire

The questionnaire assessed demographics such as gender, age, ethnicity, education level, and income. In addition, participants were asked to indicate how often they used the Internet for work or study purposes (1, [almost] never; 6, every day), for private use (1, [almost] never; 6, every day), and for finding information on health-related topics (1, never; 6, very often). For the questions assessing Internet use for work and private purposes, new categories were composed: frequent user (>3 times a week), average user (every week, but not >3 days a week), and infrequent user (<1 a week).

Furthermore, the questionnaire contained all factors identified by experts in the first round of study 1a, resulting in a questionnaire consisting of 92 questions. Potential future users were asked to rate the importance of all factors on a 7-point Likert scale ranging from 1 (not at all important) to 7 (extremely important).

Data Analysis

Data were analyzed according to the same principles used for the second round of study 1a, by calculating median scores and IQRs. Differences in consensus between the expert group and potential future users were analyzed using Wilcoxon signed rank sum tests. Similarities and differences in second-round

results between experts and potential future users were further analyzed by using multivariate analysis of variance.

Results

Study 1a: Experts

We used experts' responses from the first round to compose the questionnaire for the second round, and results of the first round are therefore shown as question items. We grouped items with regard to layout, general information content, health risk information content, ease of use, questionnaire completion, visual aids, and additional services. This resulted in a list of 85 factors thought to be related to optimal development of health-related websites. There were 10 factors in the layout category, 11 in general information content, and 16 in health risk information content. A total of 18 factors were mentioned as facilitating ease of use. The remaining factors referred to questionnaire completion (11), visual aids (6), and additional services (13). An overview of all results is given in Multimedia Appendix 1.

Consensus

During the second round, consensus was reached (IQR ≤ 1) on 57 factors. After experts rerated their answers during the third round, aided by feedback on second-round group results, consensus was obtained on another 6 factors. In total consensus was obtained on 74% of all factors.

Importance

Experts identified 33 factors as being very or extremely important (median ≥ 6): (1) layout, with 5 factors (50% of layout-related factors, eg, professional appearance and use of color), (2) general information provision, with 4 factors (36% of general information-related factors, eg, information on pros and cons of a healthy lifestyle, and tailored information on health), (3) health risk information provision, with 3 factors (19% of health risk-related factors, eg, information on skills that help to decrease health risks and personal advice on how to decrease health risks), (4) ease of use, with 10 factors (56% of factors related to ease of use, eg, availability of an easy log-in procedure and a clear navigation structure), (5) questionnaire completion, with 6 factors (55% of factors related to completion, eg, provision of a progress bar and provision of an option for partial completion), and (6) additional services, with 5 factors (38% of factors related to additional services, eg, provision of a self-monitoring tool and iterative feedback).

Although consensus was obtained on 3 factors relating to the provision of visual aids (eg, provision of cartoons, pictures, and graphical representations of relevant information), these factors had respective median scores of 4, 5, and ,5 and thus were not considered as extremely or very important. Combining the results on importance and obtained consensus, we can conclude that consensus (IQR ≤ 1) was obtained on the importance (median ≥ 6) of 24 factors (Table 3).

Table 3. Median scores for important factors in health-related Internet sites on which consensus was obtained by experts, potential users, or botha

Factor	Second round				Third round	
	Experts (n = 60)		Users (n = 120)		Experts (n = 32)	
	Median	IQR[b]	Median	IQR	Median	IQR
Which factors determine optimal layout?						
1. User friendly[c]	7	1	6	2	NA[d]	NA
2. Lively appearance[c]	6	1	5	1	NA	NA
3. Use of visual materials, such as pictures, videos, and graphics[c]	6	1	5	1	NA	NA
4. Use of color[c]	6	1	5	2	NA	NA
5. Professional appearance[e]	6	2	6	1	6	1
What kind of general information should definitely be provided on health?						
6. Information on how to attain a healthy lifestyle[c]	6	1	6	1	NA	NA
7. Information on pros and cons of a healthy lifestyle[c]	6	1	5	1	NA	NA
8. Information on health risk behaviors[c]	6	1	5	1	NA	NA
9. Personal tailored information on health[c]	6	2	5	1	6	1
What kind of health risk information should definitely be provided?						
10. Information in the form of visual aids, eg, graphs[c]	6	1	5	1	NA	NA
Which factors determine optimal ease of use?						
11. Clear structure[f]	7	1	6	1	NA	NA
12. Availability of an easy log-in procedure[f]	7	1	6	1	NA	NA
13. Use of comprehensive language[c]	7	1	6	2	NA	NA
14. Clear navigation structure[c]	7	1	2	2	NA	NA
15. Simple site design[f]	6	1	6	1	NA	NA
16. Availability of a helpdesk[c]	6	1	5	2	NA	NA
17. Availability of a function to customize the site for personal needs[c]	6	1	4	2	NA	NA
18. Availability of bread crumb navigation[e]	5	2	6	1	5	2
19. Availability of contact information for developers[c]	6	2	5	1	6	1
Which factors determine whether visitors complete questionnaires provided on the sites?						
20. Clearly structured questionnaire with clear headings and subheadings[f]	6	1	6	1	NA	NA
21. Progress bar[f]	6	1	6	1	NA	NA

Factor	Second round				Third round	
	Experts (n = 60)		Users (n = 120)		Experts (n = 32)	
	Median	IQR[b]	Median	IQR	Median	IQR
22. Opportunity to stop completion and proceed at a later time[f]	6	1	6	1	NA	NA
23. Information on personal benefits of completion[c]	6	1	5	1	NA	NA
24. Information on relevance of questionnaire completion[c]	6	1	5	1	NA	NA
25. Guarantee that completion will not result in receiving spam[e]	6	2	7	1	5	1
26. Use of original questions[e]	4	1	6	1	NA	NA
What additional services should be provided?						
27. Search engine[e]	5	1	6	1	NA	NA
28. Opportunity to print or download relevant information[c]	6	1	5	1	NA	NA
29. Opportunity for regularly revisiting the site[c]	6	1	5	1	NA	NA
30. Tool to self-monitor personal health behavior change[c]	6	1	5	2	NA	NA
31. Iterative feedback during revisits to assess users against their own previous performances[c]	6	1	5	2	NA	NA
32. Privacy statement[c]	6	2	5	2	6	1

[a] Only experts were included in the third round. Statements on which consensus was obtained in the second round were excluded from the third round and results are therefore missing. Only factors on which consensus was obtained, either by experts or users, or both, are displayed (interquartile range ≤ 1).

[b] Interquartile range.

[c] Factors on which consensus was obtained only by experts.

[d] Not applicable.

[e] Factors on which consensus was obtained only by potential users.

[f] Factors on which consensus was obtained by both experts and potential users.

Study 1b: Potential Future Users

Consensus

Potential future users reached consensus (IQR ≤ 1) on 60 of the 85 factors (71%). An overview of all results is given in Multimedia Appendix 1.

Importance

A total of 17 factors were identified as being very or extremely important (median ≥ 6). These factors were mapped into different categories: (1) layout, with 3 factors (30% of factors related to layout, eg, limited amount of distractions and user friendliness), (2) general information provision, with 1 factor (9% of factors related to general information; information on how to attain a healthy lifestyle), (3) ease of use, with 7 factors (39% of factors related to ease of use, eg, availability of an easy log-in procedure and use of comprehensive language), (4) questionnaire completion, with 3 factors (55% of factors related to questionnaire completion, eg, provision of a progress bar and provision of an option for partial completion), and (5) additional services, with 5 factors (38% of factors related to additional services, eg, provision of a self-monitoring tool and iterative feedback).

Although consensus was obtained on 14 factors relating to ease of use (eg, information on skills that help to decrease health risks, and personal advice on how to decrease health risks), all these factors had median scores of 5 and were not considered to be extremely or very important. Furthermore, all factors related to provision of visual aids had median scores <6, even though consensus was obtained on all 6 factors. Combining the results on importance and obtained consensus, we can conclude that consensus (IQR ≤ 1) was obtained on the importance (median ≥ 6) of 11 factors (Table 3).

Similarities and Differences Between Experts and Potential Future Users

The results of the second round revealed that experts agreed on the importance of 24 factors, whereas the majority of potential

future users agreed on the importance of 11 factors (median ≥ 6; IQR ≤ 1). The statistical tests showed a nonsignificant (z = –.262, P = .29) difference between the overall level of consensus obtained in the two groups.

Experts and potential future users both agreed on the importance of 7 factors: (1) general information, with 1 factor (the availability of information on how to attain a healthy lifestyle), (2) ease of use, with 3 factors (clear structure, easy log-in procedure, simple site design), and (3) questionnaire completion, with 3 factors (provision of a progress bar, a clearly structured questionnaire, and an option for partial completion).

Contrary to these similarities in perceived importance, experts and potential future users' opinions differed on 22 factors, significantly so on 19 factors (Table 4). The multivariate analysis of variance found a significant effect of group on the importance of factors related to optimal development ($F_{1,268}$ = 95.95, P < .001, R^2 = .90. Univariate F tests revealed that experts rated a lively appearance, the use of visual aids, the use of color, user friendliness (optimal layout), information in the form of visual aids (health risk information), opportunities for customizing the site, availability of a helpdesk, use of comprehensive language, simple site design, clear navigation structure (optimal ease of use), provision of relevance and personal benefits of completing the questionnaire (questionnaire completion), provision of a self-monitoring tool, option for printing and downloading information, stimulation of revisits, and iterative feedback (additional services) as significantly more important factors than did potential future users.

Conversely, potential future users rated the availability of bread crumb navigation (optimal ease of use), the provision of a guarantee that questionnaire completion will not result in receiving spam, and usage of original questions in questionnaires (questionnaire completion) as more important than did experts.

Table 4. Univariate F tests for differences in rating the importance of a factor between experts and users

Factor	Mean experts (n = 60)	Mean users (n = 210)	*F*	η^2
Optimal layout				
Lively appearance	5.48 (0.83)	5.01 (0.98)	11.37***	.04
Professional appearance	5.87 (0.97)	5.62 (0.81)	3.84	.01
Use of visual materials, such as pictures, videos, and graphics	5.98 (0.89)	5.21 (0.98)	30.23***	.10
Use of color	5.55 (0.89)	5.00 (1.05)	13.85***	.05
User friendly	6.68 (0.57)	6.09 (0.89)	24.08***	.08
General information				
Information on pros and cons of a healthy lifestyle	5.53 (0.91)	5.36 (0.97)	1.58	.01
Information on health risk behaviors	5.60 (0.81)	5.43 (0.96)	1.59	.01
Health risk information				
Information in the form of visual aids, such as graphs	5.71 (1.09)	4.88 (1.06)	28.53***	.10
Optimal ease of use				
Availability of a function to customize the site for personal needs	5.78 (0.98)	3.92 (1.36)	97.63***	.27
Availability of a helpdesk	5.68 (0.89)	5.04 (1.06)	18.09***	.06
Use of comprehensive language	6.50 (0.70)	6.01 (0.91)	14.61***	.05
Availability of bread crumb navigation	5.05 (1.19)	5.63 (0.97)	15.22***	.05
Simple site design	6.02 (0.89)	5.65 (0.89)	7.88**	.03
Clear navigation structure	6.57 (0.75)	2.18 (0.95)	1084.30**	.80
Questionnaire completion				
Information on relevance of questionnaire completion	5.77 (0.89)	5.51 (0.89)	3.92*	.01
Information on personal benefits of completion	5.78 (0.92)	5.34 (0.98)	9.95**	.04
Guarantee that completion will not result in receiving spam	5.87 (1.16)	6.29 (0.95)	8.44**	.03
Use of original questions	4.53 (1.19)	5.59 (1.01)	46.50***	.15
Additional services				
Tool to self-monitor personal health behavior change	5.88 (0.85)	4.84 (1.21)	39.41***	.13
Opportunity to print or download relevant information	6.18 (0.83)	5.41 (1.15)	23.41***	.08
Opportunity to regularly revisit the site	6.02 (0.89)	5.15 (1.13)	29.95***	.10
Iterative feedback during revisits	5.92 (0.79)	5.04 (1.07)	34.45***	.11

*$P < .05$, **$P < .01$, ***$P < .001$.

Discussion

The public health impact of Internet health communication programs is suboptimal in the target population [17,18,46]. Since program development is one of the main processes to influence reach of a new product [25], this study attempted to give both experts and users a say in the development process. By means of a Delphi study, we identified an extensive list of factors potentially related to optimal development of health-related websites, as well as the extent to which experts and users agreed on the importance of these factors. In addition, we deduced similarities and differences in perceived importance between experts and potential users.

Main Findings

This study identified an extensive list of factors that might contribute to the development of health-related websites. The importance of a selected set of these factors was stressed by both experts and potential users. Developers should therefore attempt to take these factors into account when developing a health-related website. In particular, the provision of information on attaining a healthy lifestyle was emphasized. Furthermore, a clearly structured website with a simple design and the presence of an easy log-in procedure were brought up as important factors, which corresponds to results from earlier studies [31,39]. Also, with respect to optimizing questionnaire completion, several important factors were identified. To increase questionnaire completion, websites should offer a progress bar and an option for partial completion. These results

are also in line with previous findings [47]. Since questionnaire length seems to be inversely related to actual participation as well as completion [48], providing an option for partial completion might be a solution to prevent attrition when using extensive questionnaires. Besides factors agreed upon by both groups, results also indicated that the majority of experts and users significantly differed on the importance of several factors. Due to their specific knowledge on theories in their area of expertise, experts should be involved in the development process to ensure scientific input. However, opinions of potential users must not be neglected because, by engaging users in this process, their involvement will increase. This involvement will subsequently increase the chances of obtaining higher levels of appreciation, prolonged and sustained use, and lower levels of attrition [37].

The recently introduced Internet intervention model [35] offers an opportunity to classify important factors identified in this study into several main areas that determine how the website is developed and functions. The first area pertains to the *appearance* of the website, which is one of the first website characteristics visitors are confronted with. Since previous research indicated that more than half of all website visitors are inclined to leave the website within the first 30 seconds [49], the exterior of the website should be appealing in order to attract sufficient visitor attention and prevent early disengagement. In this study, experts highlighted important factors, such as a lively appearance, the use of color, and the use of visual aids, as contributing to the appearance of health-related websites. Furthermore, potential users stressed the importance of developing websites with a professional appearance [39]. The model also emphasizes *behavioral prescription*, which refers to the instructions users receive on how to change their lifestyle. In the current study, both experts and potential users emphasized the importance of offering information on how to achieve a healthy lifestyle on health-related websites. Another important area of the website is the size of the *burden* that actual use of the website entails. Results from this study indicated that both experts and users value the presence of an easy log-in procedure. Furthermore, experts stressed the importance of developing a user-friendly website, to decrease the effort visitors must invest in navigating the website. This also entails a simple site design and a clear navigation structure. Furthermore, potential users indicated they appreciated the presence of bread crumb navigation and a search engine, to ensure visible navigation on the website [50]. *Content* that is provided on the website should be accurate, complete, and readable [2,51]. In line with these findings, experts agreed on the importance of using comprehensive language on health-related websites. Since health-related websites mainly aim at helping or assisting people to adopt a healthy lifestyle, experts stressed that detailed information on the pros and cons of a healthy lifestyle and on health behaviors in general should be included in the website's content [31,39]. Additionally, potential users indicated the importance of an opportunity to download or print relevant information that is provided on the website. *Participation* is considered as another important component of the website and focuses on its ability to engage and involve visitors. Providing reinforcement is considered an important strategy to engage visitors and can be a reward for progressing through the website content [35]. The provision of personalized feedback regarding the status of lifestyle behaviors and the opportunity to self-monitor behavior change can be regarded as a form of reinforcement. In line with this, experts stressed the importance of providing a self-monitoring tool to allow visitors to assess their current health behavior status and to track their progress. Finally, *assessment* refers to the website's ability to adjust the website to specific user needs and wishes. In this study, experts indicated that an opportunity to adjust the website to personal preferences was highly appreciated. Therefore, visitors should have an opportunity to adjust not only the appearance of the website but also its content, corresponding to their needs and wishes. Tailoring is considered to be an appropriate strategy to adjust health information to personal characteristics and preferences [52]. In addition, previous studies indicated that providing personalized, iterative feedback regarding one's health behavior might in itself stimulate revisits to the program [31]. In line with these previous findings, the majority of experts agreed on the importance of stimulating revisits and using iterative feedback. To adjust information to personal characteristics and specific user needs, questionnaires are often requisite to obtain detailed information on these topics. As stressed in the introduction, these questionnaires often tend to be extensive and therefore sensitive to early dropout. It is therefore imperative to stimulate questionnaire completion in order to optimize adjustment to visitor's needs and wishes. Experts indicated that information on relevance and personal benefit [47] of completing a questionnaire should be provided. Potential users indicated that developers should develop original questions and guarantee that questionnaire completion will not result in receiving spam.

The Internet Intervention model emphasizes two additional areas that were also identified in this Delphi study: mode of *delivery* and the *message*. Mode of delivery refers to ways in which the intervention content can be delivered to visitors. Specific strategies that can be used pertain to the use of animations, audio, video, or testimonials. The message area, on the other hand, focuses on the source and style of the message and addresses issues such as the trustworthiness and expertise of the website developers. Even though several factors pertaining to these areas were identified in the Delphi study, the majority of both experts and potential users did not agree on the importance of these factors.

Limitations

Several limitations to this study should be considered. First, the development of health-related websites is a very broad topic entailing diverse elements. To obtain detailed information on factors related to each separate element of the development process (eg, deciding on layout, content, and additional services), the questionnaire was divided into 7 categories. This rather broad setup, which is often inherent to the Delphi process, may have limited the specification of in-depth information. As the Delphi method does not allow for further specification of such factors in later rounds, still more in-depth examination of such factors is required. Second, response rates for the expert study ranged between 33% and 53%. Although this range is somewhat low, previous Delphi studies have reported similar response rates [31,43,53]. Suboptimal response rates in our

study might have been a consequence of experts being invited to participate in at least two rounds of the Delphi study. Especially, response rates among experts with a practice-based background, coming from the field of technical Web design, were low during the first round. To account for these suboptimal response rates and to balance the input from the various expert fields, we put additional effort into recruiting these experts for participation in the second round. Finally, to ensure dispersion and coverage of answers, as well as optimal participation rates, it is recommendable for future studies to limit the number of questions used in the Delphi study. Incorporating a small number of questions in the first-round questionnaire may allow experts to give more in-depth input and may increase participation in subsequent rounds, thereby diminishing attrition rates.

Implications

The currents study is a first step in increasing exposure rates of online health promotion programs. The results of this study need further experimental testing to identify which (combination of) factors ultimately results in the best result. Although the vast number of factors that play a role may hinder the feasibility of a full factorial experimental design, in-depth experimental studies on the importance of some categories of factors are recommended.

Acknowledgments

This study was funded by ZonMw, the Netherlands Organization for Health Research and Development (grant number: 120710004). Publication of the manuscript was supported by NWO, the Netherlands Organization for Scientific Research. The authors would like to thank Kim Eijmael and Claire Dolmans for their assistance during the recruitment phases and analysis of the first-round results.

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

Results of the Delphi study per factor for experts and potential future users (second and third round).

[PDF File (Adobe PDF File), 68KB - jmir_v14i1e18_app1.pdf]

References

1. Bansil P, Keenan NL, Zlot AI, Gilliland JC. Health-related information on the Web: results from the HealthStyles Survey, 2002-2003. Prev Chronic Dis 2006 Apr;3(2):A36 [FREE Full text] [Medline: 16539777]
2. Eysenbach G, Kohler C. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. AMIA Annu Symp Proc 2003:225-229. [Medline: 14728167]
3. Internet World Stats. Miniwatts Marketing Group. 2011 Oct 6. Internet Usage Statistics: The Internet Big Picture URL: http://www.internetworldstats.com/stats.htm [accessed 2011-11-25] [WebCite Cache ID 63SbuP6d0]
4. Brug J, Oenema A, Kroeze W, Raat H. The internet and nutrition education: challenges and opportunities. Eur J Clin Nutr 2005 Aug;59 Suppl 1:S130-7; discussion S138. [doi: 10.1038/sj.ejcn.1602186] [Medline: 16052181]
5. Ritterband LM, Gonder-Frederick LA, Cox DJ, Clifton AD, West RW, Borowitz S. Internet interventions: in review, in use and into the future. Prof Psychol Res Pr 2003;34:527-534. [doi: 10.1037/0735-7028.34.5.527]
6. 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] [Medline: 20558196]
7. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
8. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
9. Neville LM, O'Hara B, Milat AJ. Computer-tailored dietary behaviour change interventions: a systematic review. Health Educ Res 2009 Aug;24(4):699-720 [FREE Full text] [doi: 10.1093/her/cyp006] [Medline: 19286893]
10. Brug J, Glanz K, Van Assema P, Kok G, van Breukelen GJ. The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. Health Educ Behav 1998 Aug;25(4):517-531. [Medline: 9690108]
11. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. Patient Educ Couns 1999 Feb;36(2):107-117. [Medline: 10223016]
12. 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] [Medline: 19468084]
13. 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]

14. 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]
15. 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]
16. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Internet Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
17. 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]
18. De Nooijer JOA, Kloek G, Brug J, De Vries H, De Vries NK. Bevordering van gezondheidsgedrag via internet: nu en in de toekomst [Promotion of Healthy Behavior Through the Internet: Now and in the Future]. Maastricht: Maastricht University; 2005.
19. Crutzen R. Hard to Get, Hard to Keep: Dissemination of and Exposure to Internet-Delivered Health Behavior Change Interventions Aimed at Adolescents [dissertation]. Maastricht: Maastricht University; 2009.
20. 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]
21. Evers KE, Cummins CO, Prochaska JO, Prochaska JM. Online health behavior and disease management programs: are we ready for them? Are they ready for us? J Med Internet Res 2005 Jul 1;7(3):e27 [FREE Full text] [doi: 10.2196/jmir.7.3.e27] [Medline: 15998618]
22. Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or Internet-based surveys. Educ Psychol Meas 2000;60:821-836. [doi: 10.1177/00131640021970934]
23. 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]
24. 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]
25. Rogers E. Diffusion of Innovations. New York, NY: Free Press; 2003.
26. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res 2006;8(3):e15 [FREE Full text] [doi: 10.2196/jmir.8.3.e15] [Medline: 16954125]
27. Crutzen R, de Nooijer J, Candel MJ, de Vries NK. Adolescents who intend to change multiple health behaviours choose greater exposure to an internet-delivered intervention. J Health Psychol 2008 Oct;13(7):906-911. [doi: 10.1177/1359105308095064] [Medline: 18809641]
28. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
29. 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]
30. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345. [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
31. Brouwer W, Oenema A, Crutzen R, de Nooijer J, de Vries NK, Brug J. An exploration of factors related to dissemination of and exposure to internet-delivered behavior change interventions aimed at adults: a Delphi study approach. J Med Internet Res 2008;10(2):e10 [FREE Full text] [doi: 10.2196/jmir.956] [Medline: 18417443]
32. Moorman C, Matulich E. A model of consumers' preventive health behaviors: the role of health motivation and health ability. J Consum Res 1993;20:208-228.
33. 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] [Medline: 21212045]
34. 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] [Medline: 19632970]
35. 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]
36. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
37. Glasgow RE, Klesges LM, Dzewaltowski DA, Bull SS, Estabrooks P. The future of health behavior change research: what is needed to improve translation of research into health promotion practice? Ann Behav Med 2004 Feb;27(1):3-12. [doi: 10.1207/s15324796abm2701_2] [Medline: 14979858]

38. Cunningham M, Rapee R, Lyneham H. Feedback to a prototype self-help computer program for anxiety disorders in adolescents. Adv Ment Health 2006;5(3):216-224.
39. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
40. Adler M, Ziglio E. Gazing Into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health. Bristol, PA: Jessica Kingsley Publishers; 1996.
41. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Internet-delivered interventions aimed at adolescents: a Delphi study on dissemination and exposure. Health Educ Res 2008 Jun;23(3):427-439 [FREE Full text] [doi: 10.1093/her/cym094] [Medline: 18209115]
42. Elfeddali I, Bolman C, Mesters I, Wiers RW, de Vries H. Factors underlying smoking relapse prevention: results of an international Delphi study. Health Educ Res 2010 Dec;25(6):1008-1020 [FREE Full text] [doi: 10.1093/her/cyq053] [Medline: 20864604]
43. De Vet E, Brug J, De Nooijer J, Dijkstra A, De Vries NK. Determinants of forward stage transitions: a Delphi study. Health Educ Res 2005 Apr;20(2):195-205 [FREE Full text] [doi: 10.1093/her/cyg111] [Medline: 15328302]
44. Linstone HA, Turoff M. Introduction to the Delphi method: techniques and applications. In: Linstone HA, Turoff M, editors. The Delphi Method: Techniques and Applications. Reading, MA: Addison-Wesley; 1975:3-12.
45. FlyCatcher Internet Research. Homepage URL: http://www.flycatcher.eu/ [accessed 2011-11-25] [WebCite Cache ID 63Sbdxa2b]
46. Glasgow RE. eHealth evaluation and dissemination research. Am J Prev Med 2007 May;32(5 Suppl):S119-S126. [doi: 10.1016/j.amepre.2007.01.023] [Medline: 17466816]
47. Crutzen R. Optimalisering van de betrokkenheid van deelnemers en datakwaliteit in een cohortstudie [Optimizing Participants' Involvement and Data Quality in a Cohort Study]. Maastricht: Maastricht University; 2010.
48. Galesic M, Bosjnak M. Effects of questionnaire length on participation and indicators of response quality in a Web survey. Public Opin Q 2009;73:349-360.
49. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]
50. Bognazzini B. Nielsen Norman Group. Ask TOG: First Principles of Interaction Design URL: http://www.asktog.com/basics/firstPrinciples.html [accessed 2011-11-25] [WebCite Cache ID 63ScwBERH]
51. 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]
52. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
53. Mullen PM. Delphi: myths and reality. J Health Organ Manag 2003;17(1):37-52. [Medline: 12800279]

Abbreviations

IQR: interquartile range

Edited by G Eysenbach; submitted 07.06.11; peer-reviewed by B Cugelman; comments to author 26.10.11; revised version received 25.11.11; accepted 27.11.11; published 14.02.12.

Please cite as:
Schneider F, van Osch L, de Vries H
Identifying Factors for Optimal Development of Health-Related Websites: A Delphi Study Among Experts and Potential Future Users
J Med Internet Res 2012;14(1):e18
URL: http://www.jmir.org/2012/1/e18/
doi:10.2196/jmir.1863
PMID:22357411

Original Paper

Do Participants' Preferences for Mode of Delivery (Text, Video, or Both) Influence the Effectiveness of a Web-Based Physical Activity Intervention?

Corneel Vandelanotte[1], PhD; Mitch J Duncan[1], PhD; Ronald C Plotnikoff[2], PhD; W Kerry Mummery[3], PhD

[1]Centre for Physical Activity Studies, Institute for Health and Social Science Research, Central Queensland University, Rockhampton, Australia
[2]Priority Research Centre in Physical Activity and Nutrition, School of Education, University of Newcastle, Newcastle, Australia
[3]Faculty of Physical Education and Recreation, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Corneel Vandelanotte, PhD
Centre for Physical Activity Studies
Institute for Health and Social Science Research
Central Queensland University
Buidling 18
Bruce Highway
Rockhampton, 4701
Australia
Phone: 61 749232183
Fax: 61 749306402
Email: c.vandelanotte@cqu.edu.au

Abstract

Background: In randomized controlled trials, participants cannot choose their preferred intervention delivery mode and thus might refuse to participate or not engage fully if assigned to a nonpreferred group. This might underestimate the true effectiveness of behavior-change interventions.

Objective: To examine whether receiving interventions either matched or mismatched with participants' preferred delivery mode would influence effectiveness of a Web-based physical activity intervention.

Methods: Adults (n = 863), recruited via email, were randomly assigned to one of three intervention delivery modes (text based, video based, or combined) and received fully automated, Internet-delivered personal advice about physical activity. Personalized intervention content, based on the theory of planned behavior and stages of change concept, was identical across groups. Online, self-assessed questionnaires measuring physical activity were completed at baseline, 1 week, and 1 month. Physical activity advice acceptability and website usability were assessed at 1 week. Before randomization, participants were asked which delivery mode they preferred, to categorize them as matched or mismatched. Time spent on the website was measured throughout the intervention. We applied intention-to-treat, repeated-measures analyses of covariance to assess group differences.

Results: Attrition was high (575/863, 66.6%), though equal between groups (t_{863} =1.31, P =.19). At 1-month follow-up, 93 participants were categorized as matched and 195 as mismatched. They preferred text mode (493/803, 61.4%) over combined (216/803, 26.9%) and video modes (94/803, 11.7%). After the intervention, 20% (26/132) of matched-group participants and 34% (96/282) in the mismatched group changed their delivery mode preference. Time effects were significant for all physical activity outcomes (total physical activity: $F_{2,801}$ = 5.07, P = .009; number of activity sessions: $F_{2,801}$ = 7.52, P < .001; walking: $F_{2,801}$ = 8.32, P < .001; moderate physical activity: $F_{2,801}$ = 9.53, P < .001; and vigorous physical activity: $F_{2,801}$ = 6.04, P = .002), indicating that physical activity increased over time for both matched and mismatched groups. Matched-group participants improved physical activity outcomes slightly more than those in the mismatched group, but interaction effects were not significant. Physical activity advice acceptability (content scale: t_{368} = .10, P = .92; layout scale: t_{368} = 1.53, P = .12) and website usability (layout scale: t_{426} = .05, P = .96; ease of use scale: t_{426} = .21, P = .83) were generally high and did not differ between the matched and mismatched groups. The only significant difference (t_{621} = 2.16, P = .03) was in relation to total time spent on the website: the mismatched group spent significantly more time on the website (14.4 minutes) than the matched group (12.1 minutes).

Conclusion: Participants' preference regarding delivery mode may not significantly influence intervention outcomes. Consequently, allowing participants to choose their preferred delivery mode may not increase effectiveness of Web-based interventions.

(J Med Internet Res 2012;14(1):e37) doi:10.2196/jmir.1998

KEYWORDS

Physical activity: computer tailoring; mismatch; preferences; delivery method; website-delivered intervention; behavior-change intervention

Introduction

Physical inactivity increases the risk of developing cardiovascular disease, diabetes, hypertension, some cancers, and obesity [1,2]. As large proportions of the population are not meeting physical activity guidelines [3-5], increasing physical activity is a public health priority. As such, intervention strategies that can reach many people in a cost-effective manner are desired.

Web-based physical activity interventions have shown promising results [6-8] and will continue to gain importance through growth in Internet access (in Australia, 73% of households have broadband access), the power of Web-based applications (eg, Facebook or YouTube), and convenience (through mobile devices) [9,10]. However, the immense versatility of the Internet allows for health information to be delivered in several ways [8]. For example, interventions delivered via websites can provide personally tailored information through different modes, such as text based, video based, or both [11]. While personally tailored interventions have been shown to be effective both in offline (print based) and online studies [12-15], there are large variations in individual preferences for the mode of intervention delivery [16,17]. This raises the question as to whether the effectiveness of an intervention may be enhanced or reduced when it is provided through a preferred or nonpreferred mode [18].

This may be important, as a review of randomized controlled trials found that a substantial proportion of potential participants refused to participate in these trials for fear of being assigned to the nonpreferred option [19]. Participants may also drop out of a trial after being assigned to the nonpreferred mode, or may enter and remain in the study but not adhere to or be engaged in the treatment [20]. As such, randomly assigning participants to nonpreferred delivery modes may reduce their participation, follow-up, and satisfaction, and may thus lead to poor outcomes [21]. Conversely, allocation to the preferred intervention delivery mode may lead to greater participation and better intervention effectiveness [21]. Ideally, however, population-based interventions should be robust and optimally effective no matter which delivery mode participants prefer.

Numerous studies have compared the effectiveness of different intervention delivery modes [22-26], but few have evaluated intervention effectiveness when participants were matched with their preferred intervention delivery mode. This is because randomized controlled trials are the gold standard in intervention research, so participants do not get to choose their preferred mode of delivery [27]. Therefore, the effect that delivery mode preferences may have on intervention outcomes remains largely unknown [28]. It could be argued that randomized controlled trials underestimate the gains possible in real-life intervention implementation [29], simply because no effort is made to match delivery mode with preference.

To our knowledge, only one behavioral study has examined whether preferred and nonpreferred modes of delivery influence intervention effectiveness. In a comparison of interventions delivered by print or telephone, Lewis et al [18] found that being in the preferred group did not influence physical activity levels. A few medical studies have also examined the influence of preferences on treatment outcomes, but with inconclusive results: some studies found better effects on satisfaction [20] and effectiveness [30] when participants were allocated to the preferred group and others not [19]. Overall, there is little evidence that preference effects significantly compromise internal validity [19], but this conclusion rests on a small number of studies. If preference has indeed little impact on effectiveness, the most cost-effective practice would be to develop only the delivery mode shown to have the greatest impact on health behavior, as the development of different intervention delivery modes is time consuming and costly. This might be more important for Web-based interventions where participants are simply given a choice to receive the intervention in a different way via the website, and where developers should avoid designing costly alternative delivery modes if they will not make a substantial difference.

Nevertheless, we need to know what works for whom; as such, there is a need for algorithms that will efficiently and effectively match participants to physical activity programs that best meet their needs [31]. Computer-tailored interventions aim to do this by providing information that is as personally relevant as possible [32]. However, as alluded to above, it is perhaps not only the content of the health information that needs to be tailored to the individual to achieve optimal effectiveness [13], but also the mode by which the intervention is delivered. Therefore, this study aimed to examine the impact of receiving computer-tailored intervention content that is either matched or mismatched with participants' preferred mode of delivery (text based, video based, or both) on the acceptability, usability, and effectiveness of a Web-based physical activity intervention.

Methods

Design

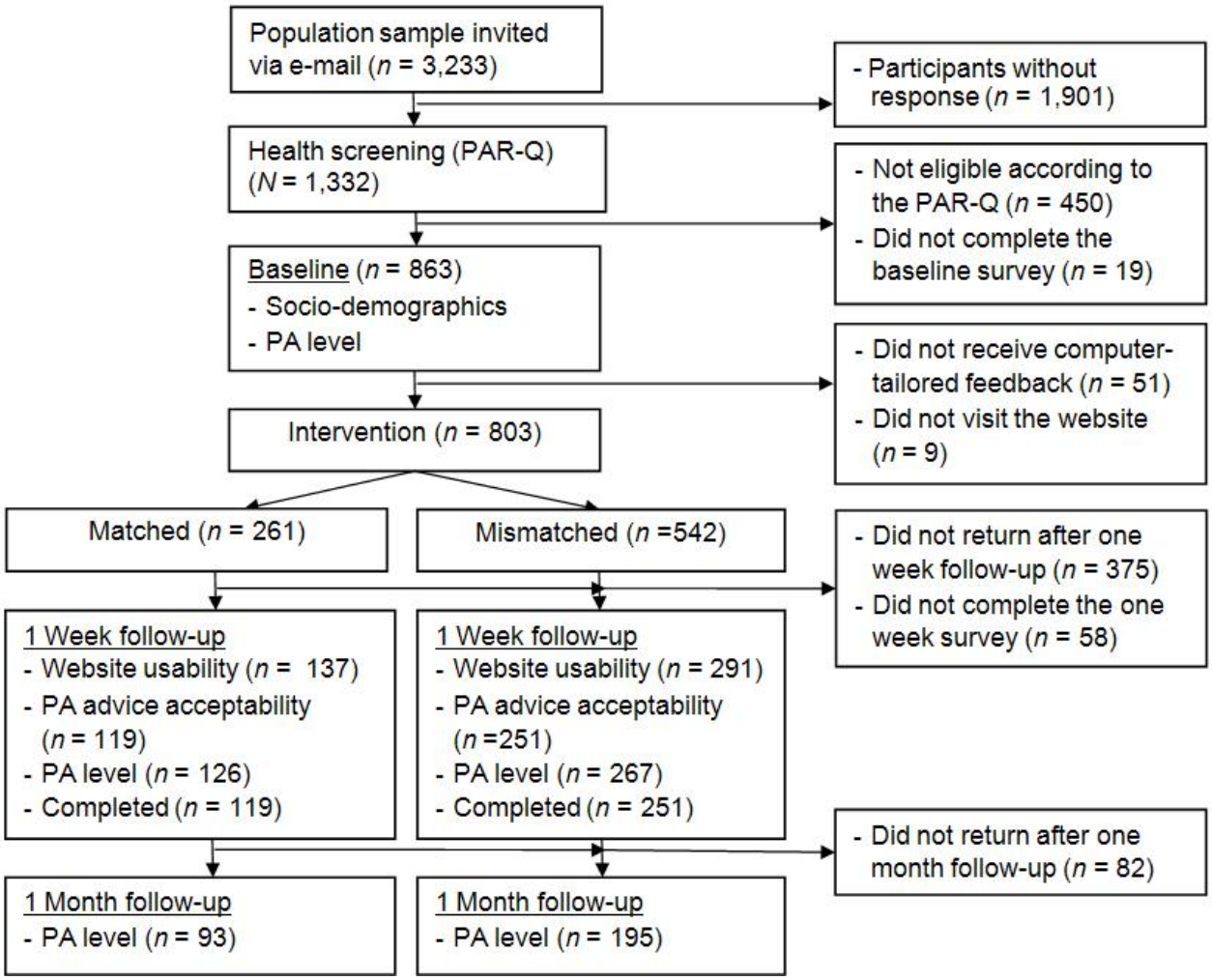
We conducted a three-arm, randomized trial with balanced allocation ratio to assess the effectiveness of a Web-based, computer-tailored physical activity intervention. The

intervention content was identical across the three arms; however, this content was delivered in a different mode for each group. The first arm received personalized physical activity feedback intervention in video mode, the second arm received the intervention in text mode, and the third arm received the intervention in combination mode, which provided both video- and text-tailored information. Data were collected during three measurement waves: baseline, 1-week follow-up, and 1-month follow-up.

Participants and Procedure

In January and February 2011, male and female adults over 18 years of age from the general population in Australia were invited by email to participate in the study. People listed in a database held by the Population Research Laboratory at the Institute of Health and Social Sciences Research at the University of Central Queensland were invited. To be eligible, participants had to have Internet access and no medical constraints that would prevent an increase in physical activity. The invitation emails contained a link to a website with information about the nature and purpose of the present study and access to the baseline survey. By accessing the baseline survey, participants provided consent to participate and agreed that they were well informed about the study. We used the Physical Activity Readiness Questionnaire (PAR-Q) to screen for participants for whom it was not safe to increase physical activity [33]. If participants answered yes to one of the PAR-Q questions, they were thanked for their time and not provided with access to the intervention website. After completing the baseline survey, participants were given a link to the intervention website; participants were automatically randomly assigned to one of the three groups on accessing the website. Nonresponders were reminded three times to complete each assessment. The whole study was entirely Web-based without any face-to-face components as part of the intervention or the assessment; as such, real-life conditions were mimicked as closely as possible. The study received ethical approval by the Human Research Ethics Committees at Central Queensland University. Figure 1 provides an overview of participant flow.

Figure 1. Participant flow. PA = physical activity; PAR-Q = Physical Activity Readiness Questionnaire.



Intervention

The intervention was based on previous Internet-delivered and computer-tailored studies that successfully increased physical activity [34-38]. However, we developed additional intervention delivery modes. In addition to the previously developed text mode, we developed a video mode and a combination mode for this study. To inform the development of the video-tailored content, focus groups and a statewide survey were conducted to explore perceived appropriateness of the new delivery modes

and volume of information presented [11]. The computer-tailored content of the three intervention modes was identical; only the intervention delivery mode was different. We did not change the intervention contents during the trial. A series of screenshots provides an impression of the intervention and shows the home page (Figure 2), an example of survey questions (Figure 3), an example of text mode (Figure 4); an example of video mode (Figure 5), and two examples of combination mode (Figure 6, Figure 7).

The intervention was largely based on the theory of planned behavior [39] and the stage of change concept [40]. Constructs of the theory of planned behavior were presented through provision of personally relevant feedback on attitudes, self-efficacy, intentions, benefits, and barriers in relation to their physical activity level. The intervention content was also adapted based on participants' stage of change, and normative feedback (whether participants met the physical activity recommendation [41]) was provided in a graph. Other nontheoretical tailored variables were age, body mass index (BMI), work environment, and the distance to often-visited places. To receive personalized physical activity advice, participants first had to complete a short questionnaire about their physical activity levels, after which the personal advice immediately appeared on screen. Participants who did not meet the physical activity recommendation were encouraged to receive more feedback by completing additional questions related to the psychosocial correlates of physical activity. Participants were provided with unlimited access to the intervention website during the intervention period.

Figure 2. Screenshot of introduction/home page.

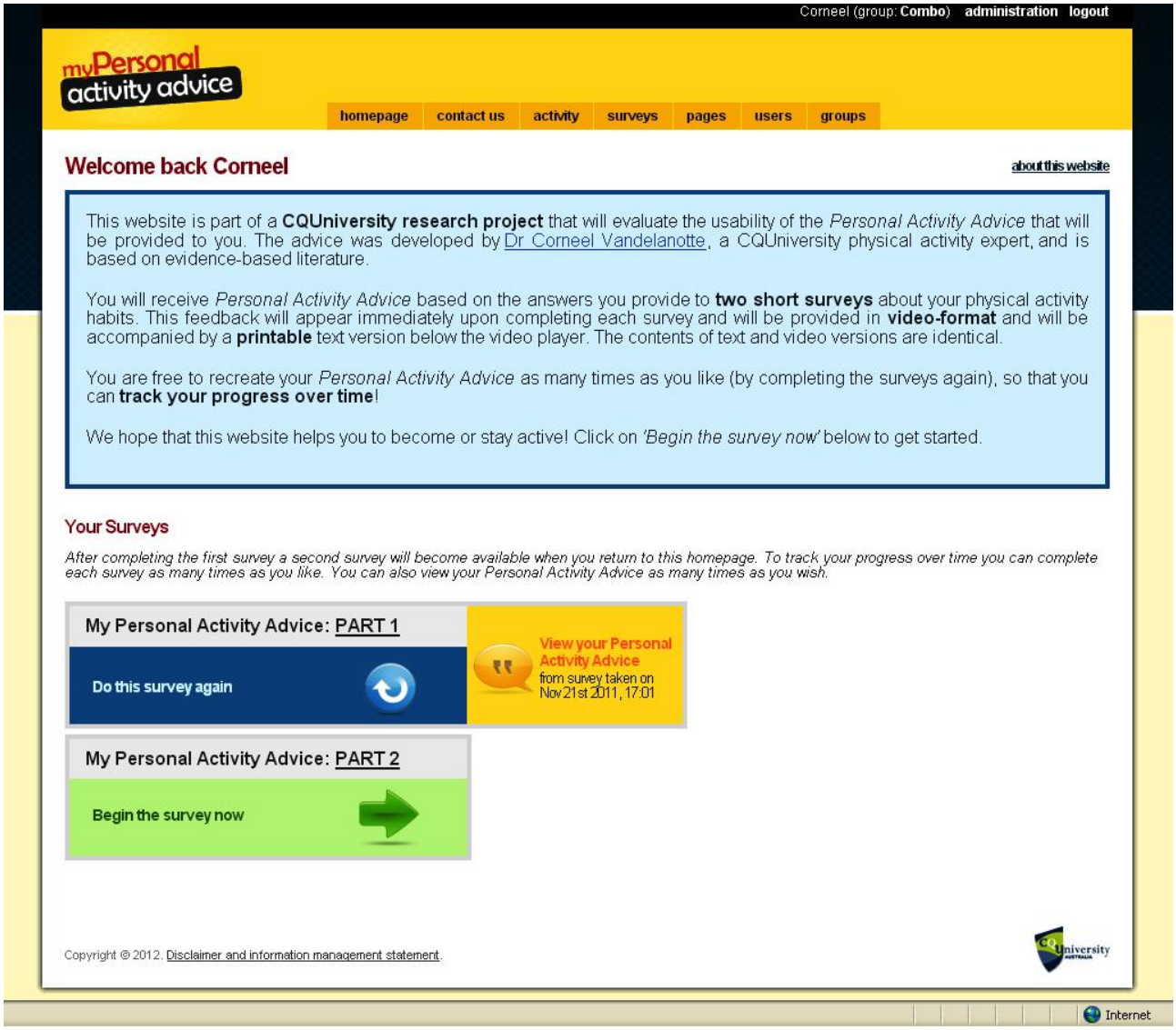


Figure 3. Screenshot of survey questions.

Figure 4. Screenshot of text mode feedback.

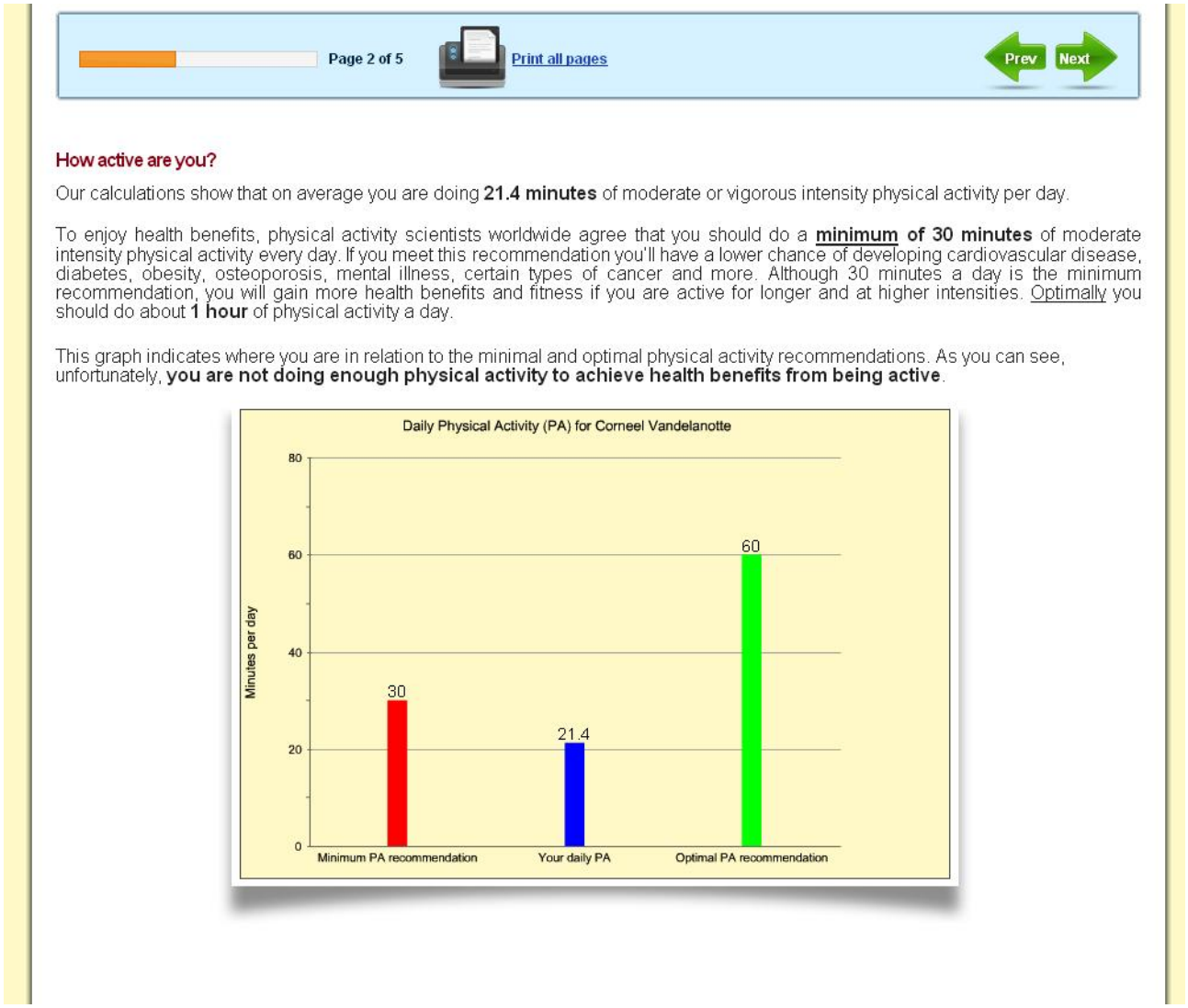


Figure 5. Screenshot of video mode feedback.

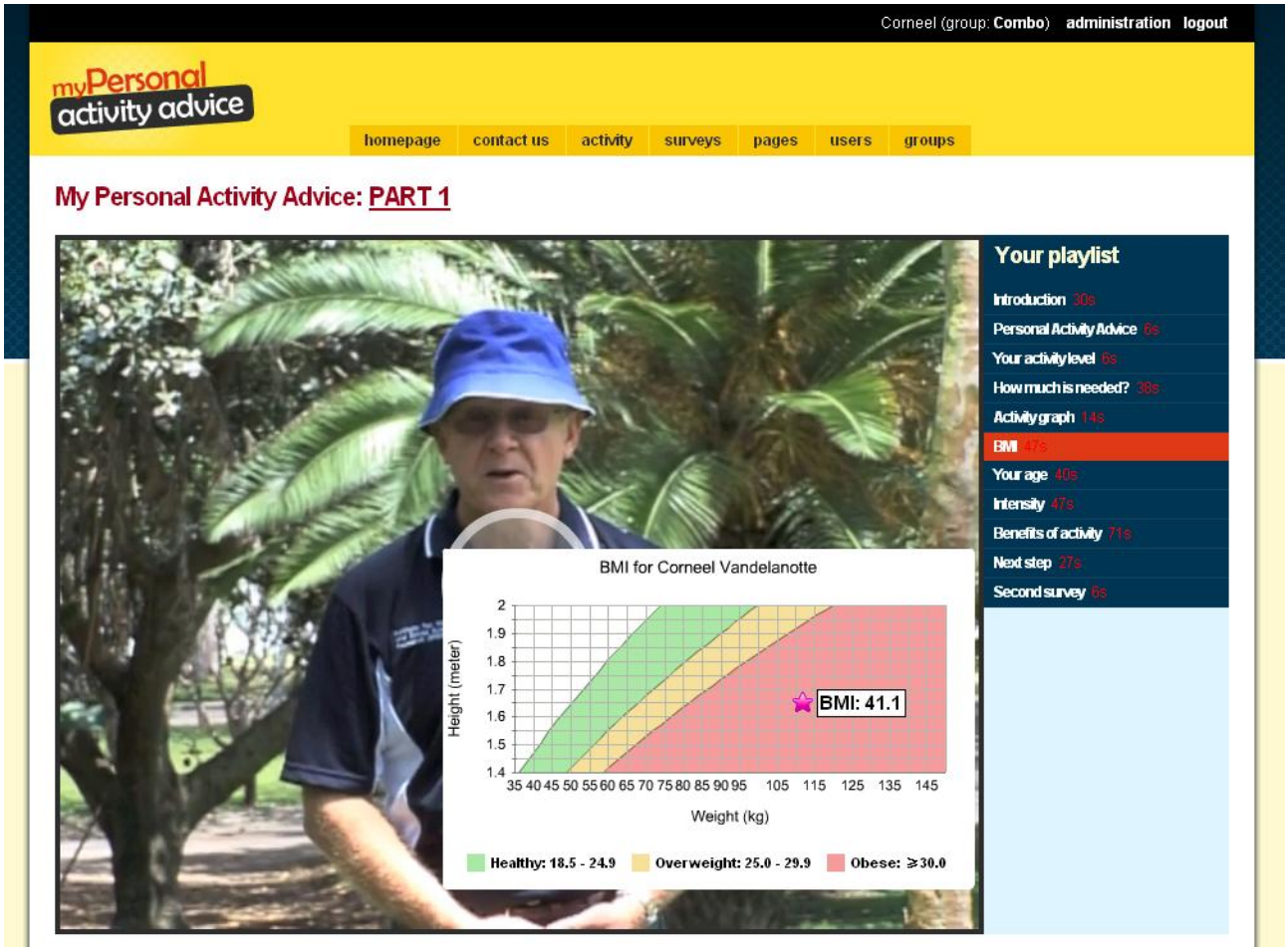


Figure 6. Screenshot of combination mode feedback 1.

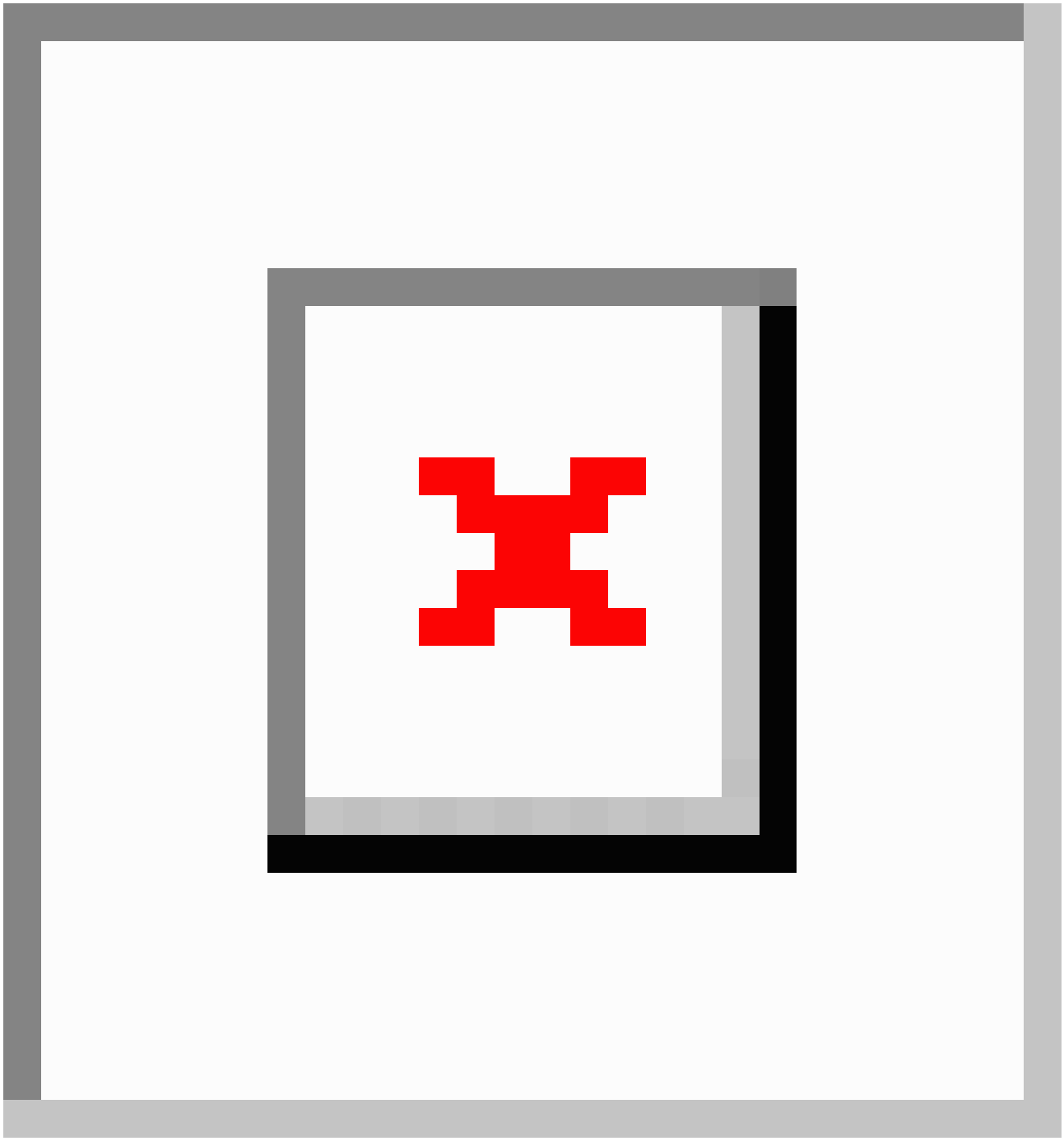
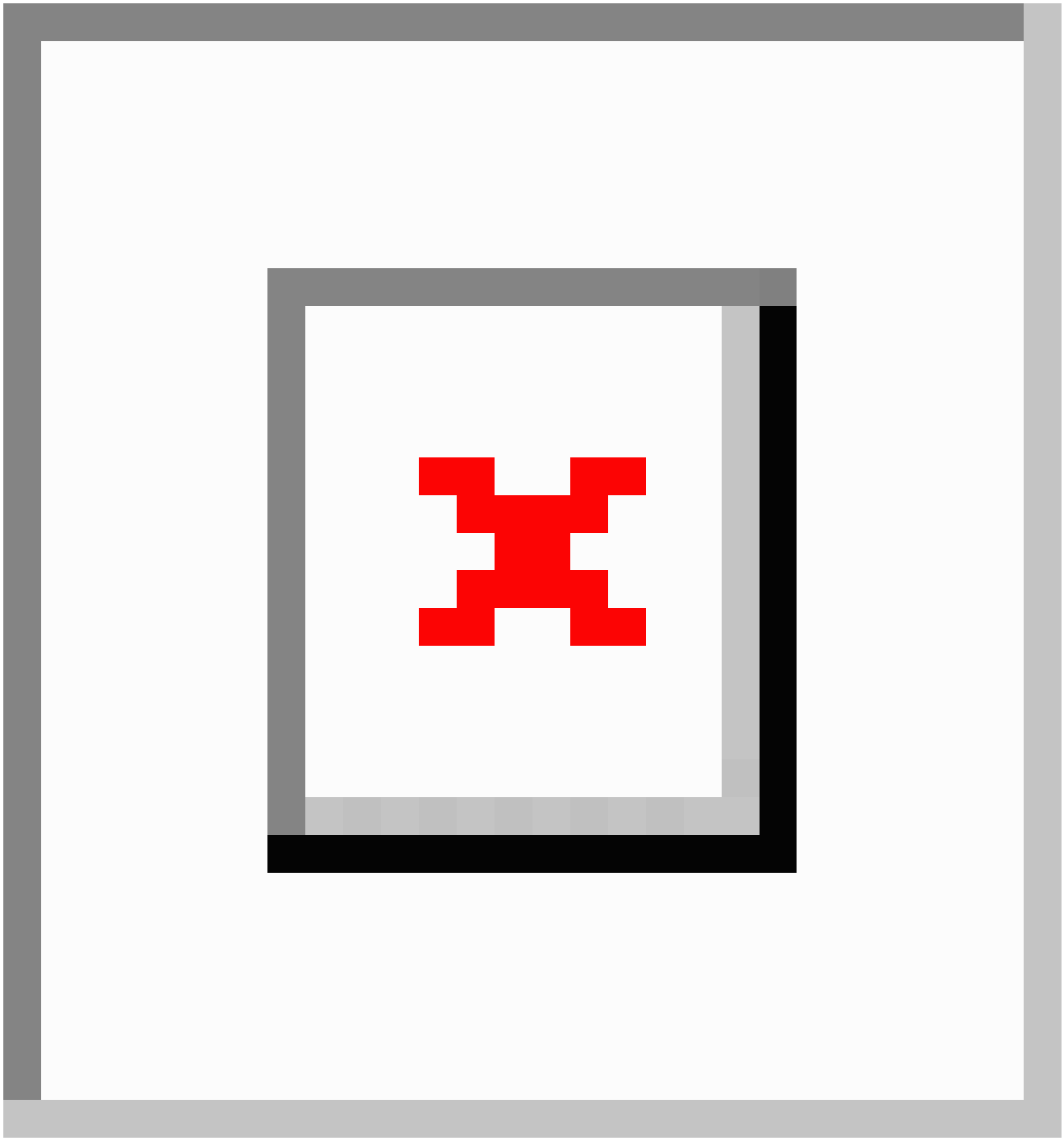


Figure 7. Screenshot of combination mode feedback 2.



Measures

We assessed all measures using a Web-based survey. The following demographic information was collected: gender, age, height and weight (to calculate BMI), employment status (unemployed; employed), level of education (low education being up to high school; high education being university or other tertiary degree), and confidence with using the Internet ("How confident are you with using the Internet for general purposes?") whereby not confident at all, not confident, and neither confident nor not confident were scored as low confidence; and whereby very confident and confident were scored as high confidence.

Physical activity level was measured using the Active Australia Survey, which has demonstrated good validity in different population groups [42,43], as well as being sensitive to change in intervention trials [44]. Questions included items on duration and frequency of walking, and of moderate- and vigorous-intensity physical activity in the previous week. To be included, all activities had to be performed continuously for at least 10 minutes at a time.

Preferred intervention delivery mode was measured using the following question: "If you were going to receive personally relevant physical activity feedback via the Internet, which intervention mode of delivery would you prefer?" The answering options were "Personally relevant written text that can be read or printed out," "A personal video message that you can watch

online or download," and "A personal video message added with transcripts that can be printed." This question was asked during the baseline assessment (before randomization), as well as 1 week after participants received the intervention, to assess whether being exposed to a nonpreferred delivery method would change participants' opinions about their preference.

For the assessment of *physical activity advice acceptability* and *website usability* (assessed only at the 1-week time point), surveys were largely based on previous published questionnaires [36,45,46]. Physical activity advice acceptability was assessed through 13 items on a 5-point Likert scale (from 1, strongly disagree, to 5, strongly agree) and divided (using factor analysis) into two scales. The first scale, physical activity advice content, was measured by 8 items, such as "The physical activity advice is credible" (alpha = .90). The second scale, physical activity advice layout, was assessed by 5 questions, such as "I liked the format through which the physical activity advice was provided" (alpha = .87). Website usability was measured by 22 items on a 5-point Likert scale (from 1, strongly disagree, to 5, strongly agree) and also divided into two scales. The website layout scale was assessed by 8 questions, such as "I liked the overall organization of the website—the links, tabs, and buttons" (alpha = .92), and the website ease of use scale was measured by another 14 questions, such as "I was able to easily find my way around the website" (alpha = .94).

Website user statistics, measuring time spent on website, were collected during the entire 1-month intervention period.

Analysis

We created two groups based on whether the intervention delivery mode to which participants were assigned was matched or mismatched with their preferred intervention delivery mode. For example, participants were mismatched if they were assigned to receive the intervention in video mode, yet they preferred to receive the intervention in text mode. Vice versa, participants were matched if, for example, they preferred to receive the intervention in combination mode (video and text based) and were actually also randomly assigned to this group.

We used 1-way analyses of variance, independent-samples *t* tests, and chi-square tests to analyze dropout, compare baseline characteristics, and examine differences in website usability, physical activity advice acceptability, and time spent on the intervention website between the matched and mismatched groups.

According to the Active Australia Survey guidelines for analysis and reporting, we computed total physical activity minutes by summing the time spent walking and on moderate- and vigorous-intensity physical activity in the past week (vigorous-intensity physical activity was weighted by two) [47]. Walking, moderate-intensity physical activity, and vigorous-intensity physical activity were added together for the total number of physical activity sessions of the past week. To account for overreporting, we truncated each activity type on the Active Australia Survey at 14 hours per week, and total activity at a maximum of 28 hours per week. A sufficient level of physical activity was defined as being active for at least 150 minutes spread across a minimum of 5 sessions each week [47].

To evaluate the intervention effects on physical activity, we conducted repeated-measures analyses of covariance with time (baseline, 1 week, 1 month) as the within-participants factor and group (matched and mismatched) as the between-participants factor, controlled for baseline differences and delivery mode to which participants were assigned. We used both an intention-to-treat analysis (last value carry forward; n = 803) and a retained-sample analysis (n = 288). All analyses were performed using SPSS version 18.0 (IBM Corporation, Somers, NY, USA). Statistical significance was set at a level of .05.

Results

Participants

Of 3233 people contacted, 1332 (41.20%) completed the PAR-Q screening. Of these, we excluded 450 (33.8%) because they were not eligible according to the screening questionnaire. Of the 863 participants who completed the baseline questionnaire, we excluded from further analysis 51 (6%) because they did not receive any tailored feedback and 9 (1%) because they did not visit the website, yielding 803 participants. The 1-week follow-up questionnaire was completed by 370 (42.9%) respondents, and 288 (33.4%) completed the 1-month follow-up (see Figure 1).

Table 1 shows participant characteristics at baseline. The average age of all respondents was 52.4 years (range 19–89 years) and the majority were women (60.7%). Levels of education (78.5%), employment (70.9%), and Internet confidence (84.3%) were high. No significant baseline differences were observed in any of the examined variables between participants in the matched and mismatched groups, except for the intervention mode to which participants were originally randomly assigned ($\chi^2_{2,801} = 151.3$, $P < .001$). This was because preferences were distributed differently for each delivery mode.

Table 1. Participant characteristics at baseline across matched and mismatched groups.

	Total study population (n = 803)	Group[a]	
		Match (n = 261)	Mismatch (n = 542)
Gender, n(%)			
Male	316 (39.4%)	103 (39.5%)	213 (39.3%)
Female	487 (60.7%)	158 (60.5%)	329 (60.7%)
Age (years), mean (SD)	52.4 (11.9)	51.6 (12.3)	52.8 (11.8)
BMI[b] (kg/m^2), mean (SD)	27.3 (6.2)	26.7 (5.3)	27.6 (6.7)
Employment status, n(%)			
Unemployed	234 (29.1%)	77 (30%)	157 (29.0%)
Employed	569 (70.9%)	184 (70.5%)	385 (71.0%)
Internet confidence, (n, %)			
Low	126 (15.7%)	37 (14%)	89 (16%)
High	677 (84.3%)	224 (85.8%)	453 (83.6%)
Education level, n (%)			
Low	173 (21.5%)	57 (22%)	116 (21.4%)
High	630 (78.5%)	204 (78.6%)	426 (78.6%)
Baseline physical activity level, n (%)			
Insufficient	352 (43.8%)	109 (41.8%)	243 (44.8%)
Sufficient	451 (56.2%)	152 (58.2%)	299 (55.2%)
Intervention condition, n (%)			
Video based	258 (32.1%)	31 (12%)	227 (41.9%)
Text based	262 (32.6%)	159 (60.9%)	103 (19.0%)
Combination	283 (35.2%)	71 (27%)	212 (39.1%)

[a] No baseline differences were observed between matched and mismatched participants, except for intervention condition ($\chi^2_{2,801}$= 151.3, P < .001).

[b] Body mass index.

Dropout

Dropout levels differed only for the age of participants. Young people were significantly more likely than older participants to drop out during the course of the study (t_{863} = 4.23, P = .000); the mean age of dropouts (50.8 years) was lower than the mean age of completers (54.3 years). No differences in dropout levels were observed for any other variables. Specifically, when comparing matched and mismatched participants, dropout was somewhat higher in the matched group (188/281, 66.9%) than in the mismatched group (387/582, 66.4%), but this difference was not significant (t_{863} = 1.31, P = .19). In relation to this, few participants took the opportunity to return to the website more than once, and there were no significant differences between matched (8/194, 4%) and mismatched (17/429, 4%) groups (t_{621} = .65, P = .51).

Preferred Intervention Delivery Mode

Table 2 shows the distributions of delivery mode preferences between matched and mismatched groups before and after the intervention. The text mode was by far the most popular preference, the video mode was the least popular preference, and the combination mode was in the middle. At baseline, delivery mode preferences were equal between matched and mismatched participants ($\chi^2_{2,801}$ = 0.03, P = .98). After receiving the intervention, participants' overall delivery mode preferences changed, although this was not statistically significant ($\chi^2_{2,412}$ = 4.1, P = .13).

Table 2. Delivery mode preferences for matched and mismatched participants pre- and postintervention.

	Video n (%)	Text n (%)	Combination n (%)	χ^2	df	*P* value
Preintervention						
Matched	31 (12%)	159 (60.9%)	71 (27%)			
Mismatched	63 (12%)	334 (61.6%)	145 (26.8%)	0.03	2,801	.98
Postintervention						
Matched	14 (11%)	92 (70%)	26 (20%)			
Mismatched	44 (16%)	168 (59.6%)	70 (25%)	4.1	2,412	.13

While the overall proportions of participants' preferences changed little after receiving the intervention, further examination of change in preference in the matched and mismatched groups revealed that change in preference appeared to be group dependent. After having received the intervention, 20% (26/132) of the participants in the matched group changed their delivery mode preference: this was 31% (4/13) for those who preferred the video mode at baseline, 3% (2/77) for those who preferred text mode, and 48% (20/42) for those who preferred the combination mode. In the mismatched group, 34% (96/282) of participants changed their delivery mode preference: this was 61% (17/28) for those who preferred the video mode at baseline, 25% (44/175) for those who preferred text mode, and 44% (35/79) for those who preferred the combination mode.

Physical Activity Changes

Table 3 reports the outcomes of the repeated-measures analyses of covariance. Significant time effects were observed for all the different physical activity outcomes (total minutes of physical activity, total number of activity sessions, walking minutes, and minutes of moderate and vigorous physical activity) according to the intention-to-treat analysis; indicating that physical activity increased over time for both matched and mismatched groups combined. This was also the case for the analysis including only participants who completed all measurements, though the *F* values are lower and not all values are significant. The participants in the matched group improved physical activity outcomes slightly more than did those in the mismatched group, but we observed no significant interaction effects for either the intention-to-treat or completer analysis. Intervention effects for participants not meeting the physical activity recommendations were also calculated; outcomes were similar to those of the total group, though time effects were stronger (the *F* values ranged between 6 and 24). However, we again noted no interaction effects (outcomes not reported in table). Finally, both the matched and the mismatched groups had an increase of 5% of participants meeting the physical activity guidelines from baseline to 1 week and then a decrease of 2% from 1 week to 1 month; there were no significant differences between groups.

Table 3. Main and interaction effects for physical activity (mean, SD) between matched and mismatched groups.

	Intention-to-treat (n = 803)				Completers (n = 288)			
	Matched (n = 261)	Mismatched (n = 542)	Time ($F_{2,801}$)	group × time ($F_{2,801}$)	Matched (n = 93)	Mismatched (n = 195)	Time ($F_{2,284}$)	group × time ($F_{2,284}$)
Total physical activity (minutes)								
Baseline	336 (352)	315 (342)			308 (301)	314 (344)		
1 week	355 (339)	320 (325)			348 (265)	322 (293)		
1 month	362 (362)	335 (345)	5.07**	0.87	355 (332)	344 (339)	2.16	0.59
Difference	+26	+20			+47	+30		
Total physical activity sessions								
Baseline	9.7 (7.1)	9.9 (9.3)			9.0 (6.1)	9.5 (7.5)		
1 week	10.2 (7.1)	10.3 (9.8)			9.6 (5.7)	10.3 (8.9)		
1 month	10.5 (7.6)	10.6 (9.4)	7.52***	0.04	10.1 (7.4)	11.2 (8.7)	5.63**	0.23
Difference	+0.8	+0.7			+1.1	+1.7		
Walking (minutes)								
Baseline	146 (158)	134 (154)			129 (146)	142 (167)		
1 week	153 (155)	139 (150)			141 (139)	143 (140)		
1 month	162 (159)	148 (159)	8.32***	0.08	165 (156)	154 (161)	4.71*	0.93
Difference	+16	+14			+36	+12		
Moderate-intensity activity (minutes)								
Baseline	54 (116)	50 (121)			44 (83)	54 (142)		
1 week	66 (125)	56 (110)			55 (102)	56 (107)		
1 month	70 (132)	59 (111)	9.53***	1.13	56 (122)	60 (111)	0.49	0.19
Difference	+16	+9			+12	+6		
Vigorous-intensity activity (minutes)								
Baseline	69 (121)	69 (130)			66 (111)	60 (105)		
1 week	76 (118)	73 (125)			75 (102)	66 (113)		
1 month	80 (125)	79 (133)	6.04**	0.26	67 (116)	66 (122)	0.64	0.22
Difference	+11	+10			+1	+6		

*P < .05, **P < .01, ***P < .001.

Physical Activity Advice Acceptability, Website Usability, and Time Spent on the Website

Physical activity advice acceptability and website usability were generally high, and differences between the matched and mismatched groups were few (Table 4). The only significant difference (t_{621} = 2.16, P = .03) was in relation to the total time spent on the website: those in the mismatched group spent significantly more time on the website (14.4 minutes) than those in the matched group (12.1 minutes). Thus, exposure to intervention materials in the mismatched group was significantly greater than in the matched group.

Table 4. Differences between matched and mismatched groups for acceptability of activity advice, website usability, and time spent on the website (mean, SD).

	Total study population (n = 428)	Matched (n = 137)	Mismatched (n = 291)	*t* test	df
Physical activity advice acceptability [a]					
Advice content	3.2 (0.7)	3.2 (0.6)	3.2 (0.7)	0.10	368
Advice layout	3.9 (0.5)	3.8 (0.5)	3.9 (0.5)	1.53	368
Website usability [a]					
Layout	3.7 (0.5)	3.7 (0.5)	3.7 (0.6)	0.05	426
Ease of use	4.0 (0.5)	4.0 (0.5)	4.0 (0.5)	0.21	426
Time spent on website (minutes)	13.7 (12.2)	12.1 (9.6)	14.4 (13.1)	2.16[b]	621

[a] On a scale ranging from 1, strongly disagree, to 5, strongly agree.

[b] $P = .31$.

Discussion

The main finding of this study was that delivery mode preference does not influence behavioral outcomes and other outcomes that are important in the effectiveness of Web-based interventions. The acceptability, usability, and effectiveness of the physical activity intervention was not significantly different for participants matched or mismatched to their preferred intervention delivery mode (video based, text based, or combination). This finding is in line with other studies that have examined preference effects for other types of interventions within different populations [18-20,48]. Only in their study on human papillomavirus testing did McCaffery et al [30] find effects on quality of life based on preferences for different interventions. The outcomes of the current study confirm the conclusion by King et al [19]: although participants may have strong intervention preferences, there is not much support for the hypothesis that preferences significantly compromise the internal validity of randomized controlled trials. The outcomes indicate that health promotion practitioners can be guided by efficacy outcomes obtained through randomized controlled trials and do not have to accommodate participant preferences with regard to intervention delivery modes in Web-based, computer-tailored physical activity interventions, as doing so would not increase intervention effectiveness.

It is difficult to explain why no differences were found, as the findings seem counterintuitive [18]. It might be that the differences between the intervention delivery modes in this study were too small to alter effectiveness. Preferences may have more impact when, for example, comparing a face-to-face intervention with a website-delivered intervention [25,26]. Different variations of the same website-delivered intervention may have been too subtle to be influential. Or a form of recruitment bias might have been at play. As all participants were reactively recruited via email, it might be that the participants liked interacting with any kind of technology or delivery mode, thus providing little opportunity for the influence of preference. This is in line with the high levels of Internet confidence reported. Alternatively, Lewis et al [18] suggested that perhaps, after being randomly assigned to a nonpreferred intervention, participants are pleasantly surprised regarding the components of the intervention and care less about being assigned to the a nonpreferred intervention. This is in line with King's [21] assumption that the stability of attitudes will affect the internal validity; a positive experience with an intervention during a trial may change negative attitudes and weaken the effect of preference on outcomes.

Though acceptability of the advice and website usability were not influenced by delivery mode preferences in this study, others have seen an impact of participant satisfaction levels [20]. Therefore, in line with Foley et al [49], we recommend offering participants a delivery mode choice if possible, as it will not harm intervention effectiveness, and might also increase participant satisfaction levels. The current study unexpectedly found that mismatched participants spend significantly more time on the website, yet this longer exposure to intervention materials did not translate to differences in outcomes. This is in contrast to previous studies that have emphasized the importance of intervention exposure to achieve behavioral change [50-52]. However, while the differences in time spent on the website were statistically significant, we do not know whether this difference in time (approximately 2.3 minutes) was meaningful enough to alter behavior. Alternatively, it might be that the extra time spent on the website was of low quality in relation to paying attention to the physical activity message. Participants might have been distracted by emails or might have visited another website simultaneously. As the online environment is extremely competitive, further research should investigate the effects on outcomes of paying attention to website-delivered interventions.

When examining the preferences themselves, we noted that strong support for the text-based delivery mode was evident, and this remained so after participants were exposed to the text mode. Also, change in preference was very low among those who preferred and received a text mode intervention ((2/77, 3%), and it was very high for those who preferred a video mode intervention but did not receive it (17/28, 60%). This is in contrast with eye-tracking research, which has shown that people don't fully read text on the Internet; instead, they scan and skim the content [53,54]. Internet-based reading behavior is

characterized by more time spent on browsing and scanning, keyword spotting, nonlinear reading, and more selective reading, while less time is spent on in-depth and concentrated reading [53,55]. Perhaps this is different when the information offered is personally relevant, as is the case with computer-tailored interventions [12]. According to the Elaboration Likelihood Model, information is processed more thoroughly when it is perceived to be personally relevant [56]. Finally, change in delivery mode preferences, after participants received the intervention, was greatest in the mismatched group (34.0% vs 19.7%). This was not unexpected, as for matched participants a change in preference is likely related to dissatisfaction with their original choice; for mismatched participants a change in preference is likely related to satisfaction with the delivery mode to which they were exposed during the intervention over their original preference.

This study has some limitations that limit its generalizability. First, the low-intensity real-life implementation of the current study (email recruitment, no face-to-face or telephone contact for the entire study) was more than likely responsible for the high attrition levels [57,58]. Yet these dropout levels are comparable with those of other website-delivered studies with similar protocols [59-61]. Second, as also mentioned by Lewis et al [18], participants were administered a forced-choice question regarding preference. In other words, even when participants didn't have a preference, or would have preferred not to use any of the delivery modes presented to them, they were forced to make a choice. In relation to this, participants might have had a different preference if they would have been able to experience each of the delivery modes beforehand; due to the innovative nature of this intervention and its delivery modes (with which the participant would have been unfamiliar), this might have indeed been the case. However, from a practical point of view, it was not possible to expose participants to the delivery mode options prior to the randomized trial. A third limitation is that the sample was relatively homogeneous (eg, mostly white and educated). The effect of preference may possibly vary across sociodemographic variables. Fourth, a larger sample size is needed to explore the effects of preference within each intervention delivery mode (video based, text based, and combination) separately; the current study was not sufficiently powered to do so. Fifth, the follow-up period in this study was short, and potentially preferences play a greater role in maintenance and use of different delivery modes over the longer term.

In conclusion, this study illustrates that the importance of preference effects in different delivery modes of an Internet-based physical activity intervention is limited. However, due to the scarcity of research in this area, more studies to investigate this research topic that can address the above-mentioned limitations are needed.

Acknowledgments

This study was funded by a Central Queensland University Merit Grant (which is part of the Research Development and Incentive Program). The Population Research Laboratory, managed by Ms Christine Hanley, was responsible for participant recruitment and data collection using Web-based surveys. Dr Vandelanotte was supported by a National Health and Medical Research Council of Australia (#519778) and National Heart Foundation of Australia (#PH 07B 3303) postdoctoral research fellowship. Professor Plotnikoff was supported by a Senior Research Fellowship from the National Health and Medical Research Council of Australia.

Conflicts of Interest

Dr Vandelanotte is the owner and codeveloper of the intervention presented in this study.

Multimedia Appendix 1

CONSORT-EHEALTH form V1.6 [62].

[PDF File (Adobe PDF File), 812KB - jmir_v14i1e37_app1.pdf]

References

1. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1423-1434. [doi: 10.1249/mss.0b013e3180616b27] [Medline: 17762377]
2. Blair SN, Morris JN. Healthy hearts--and the universal benefits of being physically active: physical activity and health. Ann Epidemiol 2009 Apr;19(4):253-256. [doi: 10.1016/j.annepidem.2009.01.019] [Medline: 19344864]
3. Vandelanotte C, Duncan MJ, Caperchione C, Hanley C, Mummery WK. Physical activity trends in Queensland (2002 to 2008): are women becoming more active than men? Aust N Z J Public Health 2010 Jun;34(3):248-254. [doi: 10.1111/j.1753-6405.2010.00521.x] [Medline: 20618264]
4. US Department of Health and Human Services. Physical Activity and Health: A Report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1996.
5. Tucker JM, Welk GJ, Beyler NK. Physical activity in U.S.: adults compliance with the Physical Activity Guidelines for Americans. Am J Prev Med 2011 Apr;40(4):454-461. [doi: 10.1016/j.amepre.2010.12.016] [Medline: 21406280]

6. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
7. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
8. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
9. Australian Bureau of Statistics. Household Use of Information Technology, Australia, 2010-11. Canberra: Australian Bureau of Statistics; 2011. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/8146.0?OpenDocument [accessed 2012-02-14] [WebCite Cache ID 65RvBDPYU]
10. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. J Telemed Telecare 2011;17(1):41-48. [doi: 10.1258/jtt.2010.100322] [Medline: 21097565]
11. Vandelanotte C, Mummery WK. Qualitative and quantitative research into the development and feasibility of a video-tailored physical activity intervention. Int J Behav Nutr Phys Act 2011;8(1):70 [FREE Full text] [doi: 10.1186/1479-5868-8-70] [Medline: 21718546]
12. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
13. 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] [Medline: 18947966]
14. Neville LM, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. Int J Behav Nutr Phys Act 2009;6:30 [FREE Full text] [doi: 10.1186/1479-5868-6-30] [Medline: 19490649]
15. 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]
16. Daley A, Stokes-Lampard H, Wilson S, Rees M, Roalfe A, Macarthur C. What women want? Exercise preferences of menopausal women. Maturitas 2011 Feb;68(2):174-178. [doi: 10.1016/j.maturitas.2010.11.011] [Medline: 21168292]
17. Rogers LQ, Courneya KS, Verhulst S, Markwell SJ, McAuley E. Factors associated with exercise counseling and program preferences among breast cancer survivors. J Phys Act Health 2008 Sep;5(5):688-705. [Medline: 18820344]
18. Lewis BA, Napolitano MA, Whiteley JA, Marcus BH. The effect of preferences for print versus telephone interventions on compliance and attrition in a randomized controlled physical activity trial. Psychol Sport Exerc 2006;7(5):453-462. [doi: 10.1016/j.psycholsport.2006.04.004]
19. King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, et al. Impact of participant and physician intervention preferences on randomized trials: a systematic review. JAMA 2005 Mar 2;293(9):1089-1099 [FREE Full text] [doi: 10.1001/jama.293.9.1089] [Medline: 15741531]
20. Floyd AH, Moyer A. Effects of participant preferences in unblinded randomized controlled trials. J Empir Res Hum Res Ethics 2010 Jun;5(2):81-93. [doi: 10.1525/jer.2010.5.2.81] [Medline: 20569152]
21. King M. The effects of patients' and practitioners' preferences on randomized clinical trials. Palliat Med 2000 Nov;14(6):539-542. [Medline: 11219886]
22. Marcus BH, Lewis BA, Williams DM, Dunsiger S, Jakicic JM, Whiteley JA, et al. A comparison of Internet and print-based physical activity interventions. Arch Intern Med 2007 May 14;167(9):944-949 [FREE Full text] [doi: 10.1001/archinte.167.9.944] [Medline: 17502536]
23. Marcus BH, Napolitano MA, King AC, Lewis BA, Whiteley JA, Albrecht A, et al. Telephone versus print delivery of an individualized motivationally tailored physical activity intervention: Project STRIDE. Health Psychol 2007 Jul;26(4):401-409. [doi: 10.1037/0278-6133.26.4.401] [Medline: 17605559]
24. Steele R, Mummery WK, Dwyer T. Using the Internet to promote physical activity: a randomized trial of intervention delivery modes. J Phys Act Health 2007 Jul;4(3):245-260. [Medline: 17846455]
25. Steele RM, Mummery WK, Dwyer T. A comparison of face-to-face or internet-delivered physical activity intervention on targeted determinants. Health Educ Behav 2009 Dec;36(6):1051-1064. [doi: 10.1177/1090198109335802] [Medline: 19502534]
26. Steele RM, Mummery WK, Dwyer T. Examination of program exposure across intervention delivery modes: face-to-face versus internet. Int J Behav Nutr Phys Act 2007;4:7 [FREE Full text] [doi: 10.1186/1479-5868-4-7] [Medline: 17352817]
27. Bonell CP, Hargreaves J, Cousens S, Ross D, Hayes R, Petticrew M, et al. Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions. J Epidemiol Community Health 2011 Jul;65(7):582-587. [doi: 10.1136/jech.2008.082602] [Medline: 19213758]
28. Torgerson DJ, Klaber-Moffett J, Russell IT. Patient preferences in randomised trials: threat or opportunity? J Health Serv Res Policy 1996 Oct;1(4):194-197. [Medline: 10180870]

29. Udell JA, Redelmeier DA. Patient preferences and the ironic nature of randomized trials. Med Decis Making 2011;31(2):226-228. [doi: 10.1177/0272989X11399125] [Medline: 21393562]
30. McCaffery KJ, Turner R, Macaskill P, Walter SD, Chan SF, Irwig L. Determining the impact of informed choice: separating treatment effects from the effects of choice and selection in randomized trials. Med Decis Making 2011;31(2):229-236. [doi: 10.1177/0272989X10379919] [Medline: 21041538]
31. King AC, Marcus B, Ahn D, Dunn AL, Rejeski WJ, Sallis JF, Activity Counseling Trial Research Group. Identifying subgroups that succeed or fail with three levels of physical activity intervention: the Activity Counseling Trial. Health Psychol 2006 May;25(3):336-347. [doi: 10.1037/0278-6133.25.3.336] [Medline: 16719605]
32. Vandelanotte C, De Bourdeaudhuij I, Sallis JF, Spittaels H, Brug J. Efficacy of sequential or simultaneous interactive computer-tailored interventions for increasing physical activity and decreasing fat intake. Ann Behav Med 2005 Apr;29(2):138-146. [doi: 10.1207/s15324796abm2902_8] [Medline: 15823787]
33. Cardinal BJ, Esters J, Cardinal MK. Evaluation of the revised physical activity readiness questionnaire in older adults. Med Sci Sports Exerc 1996 Apr;28(4):468-472. [Medline: 8778552]
34. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396 [FREE Full text] [doi: 10.1093/her/cyl096] [Medline: 16971674]
35. Spittaels H, De Bourdeaudhuij I, Vandelanotte C. Evaluation of a website-delivered computer-tailored intervention for increasing physical activity in the general population. Prev Med 2007 Mar;44(3):209-217. [doi: 10.1016/j.ypmed.2006.11.010] [Medline: 17197015]
36. Vandelanotte C, De Bourdeaudhuij I. Acceptability and feasibility of a computer-tailored physical activity intervention using stages of change: project FAITH. Health Educ Res 2003 Jun;18(3):304-317 [FREE Full text] [Medline: 12828232]
37. Vandelanotte C, De Bourdeaudhuij I, Brug J. Two-year follow-up of sequential and simultaneous interactive computer-tailored interventions for increasing physical activity and decreasing fat intake. Ann Behav Med 2007 Apr;33(2):213-219. [doi: 10.1080/08836610701310086] [Medline: 17447874]
38. Vandelanotte C, Reeves MM, Brug J, De Bourdeaudhuij I. A randomized trial of sequential and simultaneous multiple behavior change interventions for physical activity and fat intake. Prev Med 2008 Mar;46(3):232-237. [doi: 10.1016/j.ypmed.2007.07.008] [Medline: 17707079]
39. 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]
40. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
41. Department of Health and Aged Care (DHAC). National Physical Activity Guidelines for Australians. Canberra: Australian Government, Department of Health and Ageing; 1999.
42. Brown WJ, Trost SG, Bauman A, Mummery K, Owen N. Test-retest reliability of four physical activity measures used in population surveys. J Sci Med Sport 2004 Jun;7(2):205-215. [Medline: 15362316]
43. Heesch KC, Hill RL, van Uffelen JG, Brown WJ. Are Active Australia physical activity questions valid for older adults? J Sci Med Sport 2011 May;14(3):233-237. [doi: 10.1016/j.jsams.2010.11.004] [Medline: 21276752]
44. Reeves MM, Marshall AL, Owen N, Winkler EA, Eakin EG. Measuring physical activity change in broad-reach intervention trials. J Phys Act Health 2010 Mar;7(2):194-202. [Medline: 20484758]
45. Vandelanotte C, De Bourdeaudhuij I, Brug J. Acceptability and feasibility of an interactive computer-tailored fat intake intervention in Belgium. Health Promot Int 2004 Dec;19(4):463-470 [FREE Full text] [doi: 10.1093/heapro/dah408] [Medline: 15522947]
46. Hinchliffe A, Mummery WK. Applying usability testing techniques to improve a health promotion website. Health Promot J Austr 2008 Apr;19(1):29-35. [Medline: 18481929]
47. Australian Institute of Health and Welfare. The Active Australia Survey: A Guide and Manual for Implementation, Analysis and Reporting. Canberra: Australian Institute of Health and Welfare; 2003.
48. Ward J, Wilkinson C, Graser SV, Prusak KA. Effects of choice on student motivation and physical activity behaviour in physical education. J Teach Phys Educ 2008;27:385-398.
49. Foley L, Maddison R, Jones Z, Brown P, Davys A. Comparison of two modes of delivery of an exercise prescription scheme. N Z Med J 2011 Jul 8;124(1338):44-54. [Medline: 21946962]
50. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 steps australia. J Med Internet Res 2012;14(1):e4 [FREE Full text] [doi: 10.2196/jmir.1792] [Medline: 22260810]
51. van den Berg MH, Ronday HK, Peeters AJ, Voogt-van der Harst EM, Munneke M, Breedveld FC, et al. Engagement and satisfaction with an Internet-based physical activity intervention in patients with rheumatoid arthritis. Rheumatology (Oxford) 2007 Mar;46(3):545-552 [FREE Full text] [doi: 10.1093/rheumatology/kel341] [Medline: 17043050]
52. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer 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]

53. Liu Z. Reading behavior in the digital environment: changes in reading behavior over the past ten years. J Doc 2005;61(6):700-712. [doi: 10.1108/00220410510632040]
54. Nielsen J, Pernice K. Eyetracking Web Usability (Voices That Matter). Berkeley, CA: New Riders Press; 2010.
55. Sutherland-Smith W. Weaving the literacy Web: changes in reading from page to screen. Read Teach 2002;55:662-669.
56. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19:123-205. [doi: 10.1016/S0065-2601(08)60214-2]
57. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr 29;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
58. 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]
59. 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]
60. Kosma M, Cardinal BJ, McCubbin JA. A pilot study of a web-based physical activity motivational program for adults with physical disabilities. Disabil Rehabil 2005 Dec 15;27(23):1435-1442. [doi: 10.1080/09638280500242713] [Medline: 16418058]
61. Woolf SH, Krist AH, Johnson RE, Wilson DB, Rothemich SF, Norman GJ, et al. A practice-sponsored Web site to help patients pursue healthy behaviors: an ACORN study. Ann Fam Med 2006;4(2):148-152 [FREE Full text] [doi: 10.1370/afm.522] [Medline: 16569718]
62. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

BMI: body mass index
PAR-Q: Physical Activity Readiness Questionnaire

Edited by G Eysenbach; submitted 16.11.11; peer-reviewed by S Sahlqvist; comments to author 07.12.11; revised version received 29.01.12; accepted 06.02.12; published 29.02.12.

Please cite as:
Vandelanotte C, Duncan MJ, Plotnikoff RC, Mummery WK
Do Participants' Preferences for Mode of Delivery (Text, Video, or Both) Influence the Effectiveness of a Web-Based Physical Activity Intervention?
J Med Internet Res 2012;14(1):e37
URL: http://www.jmir.org/2012/1/e37/
doi:10.2196/jmir.1998
PMID:22377834

Original Paper

An Internet-Based Virtual Coach to Promote Physical Activity Adherence in Overweight Adults: Randomized Controlled Trial

Alice Watson[1,2], MBChB, MRCP, MPH; Timothy Bickmore[3], PhD; Abby Cange[1], BS; Ambar Kulshreshtha[1,2], MD, MPH; Joseph Kvedar[1,2], MD

[1]Center for Connected Health, Partners HealthCare, Boston, MA, United States

[2]Harvard Medical School, Boston, MA, United States

[3]College of Computer and Information Science, Northeastern University, Boston, MA, United States

Corresponding Author:
Alice Watson, MBChB, MRCP, MPH
Center for Connected Health
Partners HealthCare
25 New Chardon Street
Suite 400D
Boston, MA, 02114
United States
Phone: 1 617 913 2620
Fax: 1 617 726 7530
Email: ajwatson@partners.org

Abstract

Background: Addressing the obesity epidemic requires the development of effective, scalable interventions. Pedometers and Web-based programs are beneficial in increasing activity levels but might be enhanced by the addition of nonhuman coaching.

Objectives: We hypothesized that a virtual coach would increase activity levels, via step count, in overweight or obese individuals beyond the effect observed using a pedometer and website alone.

Methods: We recruited 70 participants with a body mass index (BMI) between 25 and 35 kg/m^2 from the Boston metropolitan area. Participants were assigned to one of two study arms and asked to wear a pedometer and access a website to view step counts. Intervention participants also met with a virtual coach, an automated, animated computer agent that ran on their home computers, set goals, and provided personalized feedback. Data were collected and analyzed in 2008. The primary outcome measure was change in activity level (percentage change in step count) over the 12-week study, split into four 3-week time periods. Major secondary outcomes were change in BMI and participants' satisfaction.

Results: The mean age of participants was 42 years; the majority of participants were female (59/70, 84%), white (53/70, 76%), and college educated (68/70, 97%). Of the initial 70 participants, 62 completed the study. Step counts were maintained in intervention participants but declined in controls. The percentage change in step count between those in the intervention and control arms, from the start to the end, did not reach the threshold for significance (2.9% vs –12.8% respectively, P = .07). However, repeated measures analysis showed a significant difference when comparing percentage changes in step counts between control and intervention participants over all time points (analysis of variance, P = .02). There were no significant changes in secondary outcome measures.

Conclusions: The virtual coach was beneficial in maintaining activity level. The long-term benefits and additional applications of this technology warrant further study.

Trial Registration: ClinicalTrials.gov NCT00792207; http://clinicaltrials.gov/ct2/show/NCT00792207 (Archived by WebCite at http://www.webcitation.org/63sm9mXUD)

(J Med Internet Res 2012;14(1):e1) doi:10.2196/jmir.1629

KEYWORDS

Activity monitoring; pedometers; obesity; body mass index; telemedicine; telehealth

Introduction

With 65% of US adults being overweight, and a third meeting the criteria for obesity [1], health professionals have been spurred to develop innovative strategies to address this epidemic. A major driver of obesity is inadequate physical activity [2], with only a quarter of US adults engaging in the recommended amount of weekly physical activity [3].

Motivational coaching, personalized feedback, goal setting, and patient education have been used successfully to bring about long-term changes in diet and activity [4,5]. Delivering these components via traditional means, however, such as face-to-face interactions with a clinician, is costly and difficult to scale up.

New technologies have shown promise as effective, accessible, and inexpensive solutions. Pedometers, wearable devices that capture step count, can increase activity levels by up to 2000 steps per day [6]. In addition, the use of the Internet to communicate information to individuals may allow population health interventions to be delivered, and widely disseminated, at relatively low cost [7-10].To date, several Internet-based interventions have demonstrated reductions in weight through a combination of self-monitoring, education, and motivational messaging [11-13].

Such interventions may be enhanced by the addition of a coach, to promote accountability and adherence to an exercise program. In the helping professions there is a well-documented association between the quality of the professional–client relationship and outcomes [14]. The psychotherapy literature describes this working alliance as the trust and belief that the helper and patient have in each other as team members in achieving a desired outcome [15,16].

Coaching need not be carried out face-to-face; indeed, formal and informal e-coaching models have been demonstrated to be beneficial in promoting activity and weight loss [17,18]. Furthermore, a coach need not be human; recent research into the use of embodied computer agents has shown that participants can successfully form a working alliance relationship with a nonhuman agent or, simply put, a virtual coach [19].

In our study we sought to understand the effectiveness of virtual coaching compared with the use of a pedometer and website alone in improving activity levels in overweight or obese participants. We hypothesized that use of a virtual coach would increase their activity levels, in the form of step count, beyond the effect observed using a pedometer and website alone.

Methods

Eligibility Criteria

Participants were between 20 and 55 years old (inclusive); had a body mass index (BMI) between 25 and 35 kg/m^2 (inclusive); were fluent in spoken and written English; had a primary care physician; had access to a personal computer with an available USB port, speakers, and Internet access; and either answered no to all 7 questions on the Physical Activity Readiness Questionnaire (PAR-Q) or obtained written permission from their primary care physician to take part in the study. The PAR-Q is a validated screening tool designed to identify adults with medical problems that might preclude them from safely initiating an exercise regimen [20].

Setting

Recruitment took place in Boston, Massachusetts, USA through advertisements in local newspapers, on a local website (Craigslist), at health care facilities, and through broadcast emails within the hospital email network. All study visits took place at Massachusetts General Hospital, Boston.

Recruitment commenced in June 2008 and took 3 weeks. Data were collected and analyzed in 2008. The study was reviewed and approved in July 2007 by the institutional review board of the Massachusetts General Hospital. This trial was registered on clinicaltrials.gov (NCT00792207).

Interventions

At an initial visit, all potential participants were weighed and measured using a waist-high, stand-on bariatric scale (Detecto, capacity 600 × 0.2 lb; Cardinal Scale Manufacturing Company, Webb City, MO, USA) and a mechanical wall-mounted stadiometer, both maintained by Massachusetts General Hospital Biomedical Engineering. Following confirmation of eligibility and consent, the participant was randomly assigned to the intervention or control arm of the study.

All participants were provided with the pedometer (ActiPed; FitLinxx, Norwalk, CT, USA) and instructed to wear it at all times over the 12-week study period, apart from when bathing or sleeping. The ActiPed device is a highly accurate activity monitor, worn on the shoe, which contains an accelerometer and tracks steps [21]. The device wirelessly transmits activity data to a USB receiver on a participant's desktop computer, where it is then relayed over the Internet to a database on a secure computer server (Figure 1). All participants were given access to the pedometer manufacturer's password-protected ActiHealth website to view graphs of their activity levels over time and set personal goals.

In addition, those in the intervention arm of the study were provided access to the virtual coach, a computer-animated exercise advisor that runs using software installed on users' home computers. The virtual coach is entirely automated and follows an algorithm-driven script, using simulated face-to-face conversation, including verbal and nonverbal relationship-building behaviors modeled on best practices from studies of patient–provider health communication with the goal of establishing a working alliance [19]. The scripts used by the virtual coach were developed through an interdisciplinary collaboration involving physicians, computer scientists, and exercise trainers to ensure adherence to best practices. The script employs behavioral and social cognitive strategies demonstrated in the literature to promote exercise behavior change. These strategies include goal setting, shaping, self-monitoring, positive reinforcement, problem solving, education, and social support [22,23].

The virtual coach software was integrated with the database containing the participants' activity data to allow tailored interactions according to each participant's adherence to step

count goals. The interactions all followed a structured pattern, starting with greeting and social interaction, proceeding to review of pedometer step count, feedback and goal setting, tips on activity or diet, and commitment to date of next interaction, and ending with encouragement and farewell. However, both the dialogue structure and the format and content of individual utterances were tailored based on each user's progress in the system (eg, whether they had progressed past baseline), their current status (eg, whether they had met their short-term goals), and discourse context (eg, whether they had just asked the virtual coach a question or asked for help). As a result, those who had not met their activity target would have a different interaction at the same time point in the study from those who had met their goals. Users had to select from a series of answer options, as the system was not designed to handle free-text responses.

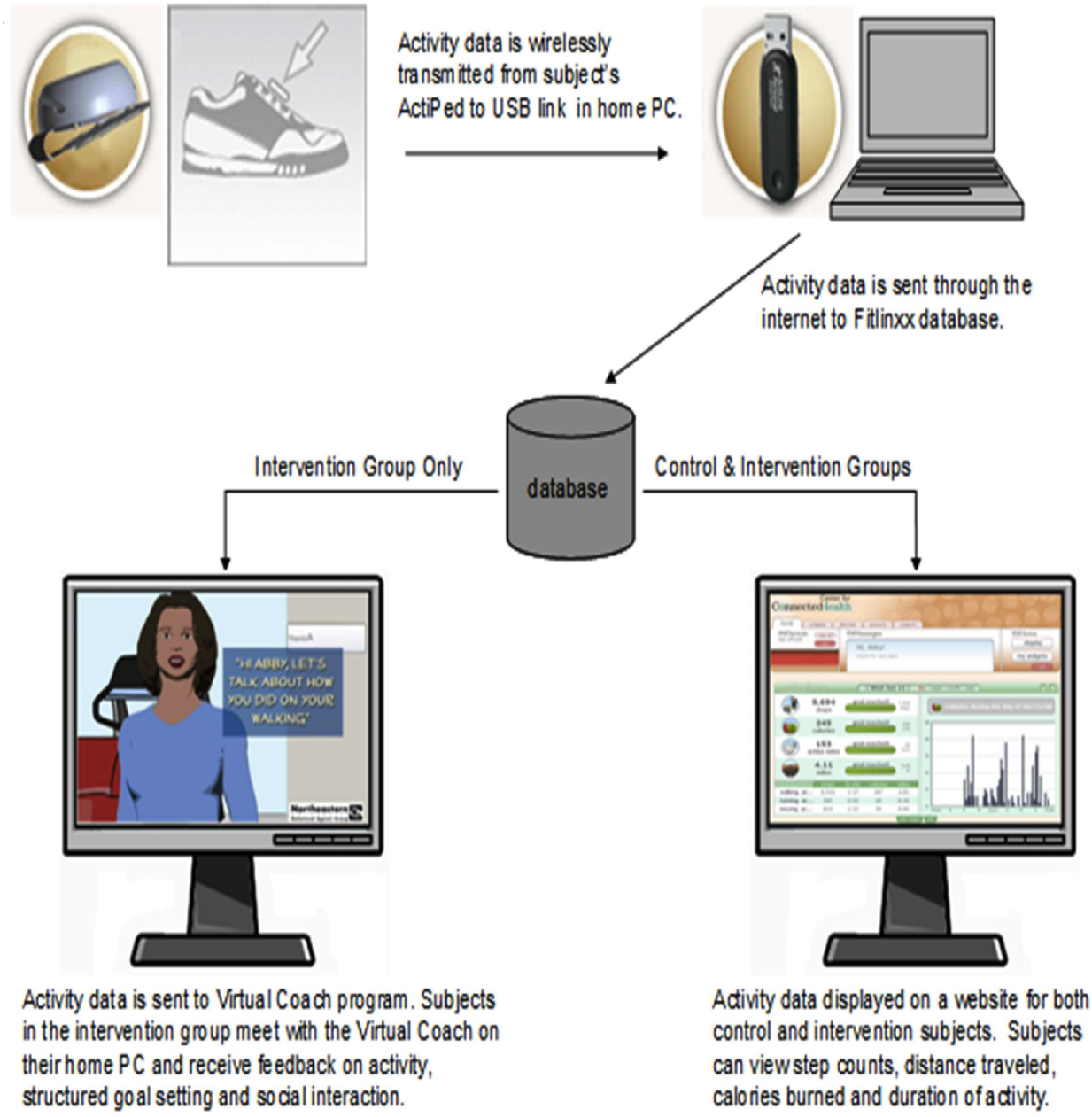
Software modules were significantly tested by the development team, followed by several end-to-end pilot tests of the intervention prior to deployment.

Intervention participants were instructed to meet with the coach three times a week throughout the study. These interactions lasted approximately 5 to 10 minutes per session. The 12-week program focused on building rapport and establishing baseline activity levels, followed by tips to increase activity, daily personalized goal setting, and advice about maintaining a healthy diet and activity level after the study concludes.

To ensure compliance and minimize loss of data throughout the 12-week study period, study staff contacted participants either if no step data were received or if those in the intervention arm did not talk to the virtual coach for 7 consecutive days. At the final visit, participants were weighed and measured and asked to complete several surveys.

Participants were provided with gift cards for attendance at each study visit. The technology was provided at no charge, although all participants were required to have a computer with Internet access.

Figure 1. Illustration of technology used in the study and flow of data.



Primary and Secondary Outcomes

The primary end point of this study was the percentage change in step count between those in the intervention and those in the control arms, from the start of the study compared with the end of the study. To evaluate the percentage change in step count, we divided the 12-week study into four 3-week time periods (P1, P2, P3, P4). Secondary end points were changes in weight, BMI, 7-day physical activity recall, physical activity stage of change, self-efficacy and exercise benefits and barriers, and satisfaction with the program [24-27]. Satisfaction was measured using a combination of novel questions regarding the activity monitor and standardized questions from the Working Alliance Inventory to assess the strength of social bond between intervention participants and the virtual coach [16].

Sample Size

This study was designed to detect a difference of 12 in the percentage change in step count between intervention and control participants from P1 to P4. This number was determined following a priori discussions with clinical experts and was thought to represent a clinically significant change in steps assuming a baseline of between 5000 and 7000 steps per day. To have an 80% power to detect such a difference, assuming a standard deviation of 16, a type I error level of .05, and a dropout rate of 20%, a total sample size of 70 was required.

Randomization

Random numbers were generated (using Microsoft Excel; Microsoft Corporation, Redmond, WA, USA) and assigned control or intervention status on a 1:1 basis with a block size of four. Sealed, ordered envelopes were prepared by one investigator, who was not involved in participant enrollment,

containing information about group assignment and opened only after the participant consented. Due to the nature of the intervention, study staff and participants were not blinded to group assignment over the course of the study.

Statistical Analysis

We compared continuous outcomes between groups using a Student *t* test (for normally distributed outcomes) and the Wilcoxon rank sum test (for nonnormally distributed outcomes and rank measures). Differences in proportions between groups were compared by using chi-square tests or Fisher exact test when appropriate. All calculations were performed with SAS version 9.1 (The SAS system for Windows; SAS Institute, Cary, NC, USA). We calculated average step counts for each 3-week period of time (period) by dividing the total number of steps recorded in the period by the number of days data were received. Days with a recorded step count of <100 were noted as missing data for the day, as this low level of activity was more likely to reflect the ActiPed being carried in a bag or moved in a house than actually being worn. Participants with no data for a period were noted as missing for this period. The analysis was conducted both examining only those participants with data for each period and including those who had missing data points in one, or more than one, period. A 2-sided *P* value of .05 was considered statistically significant. Average values are represented as mean (SE) unless otherwise stated.

Results

Patient Demographics

A total of 70 participants were enrolled, of whom 62 (89%) completed the study. The final participant completed the study in September 2008. Further details regarding enrollment are provided in Figure 2.

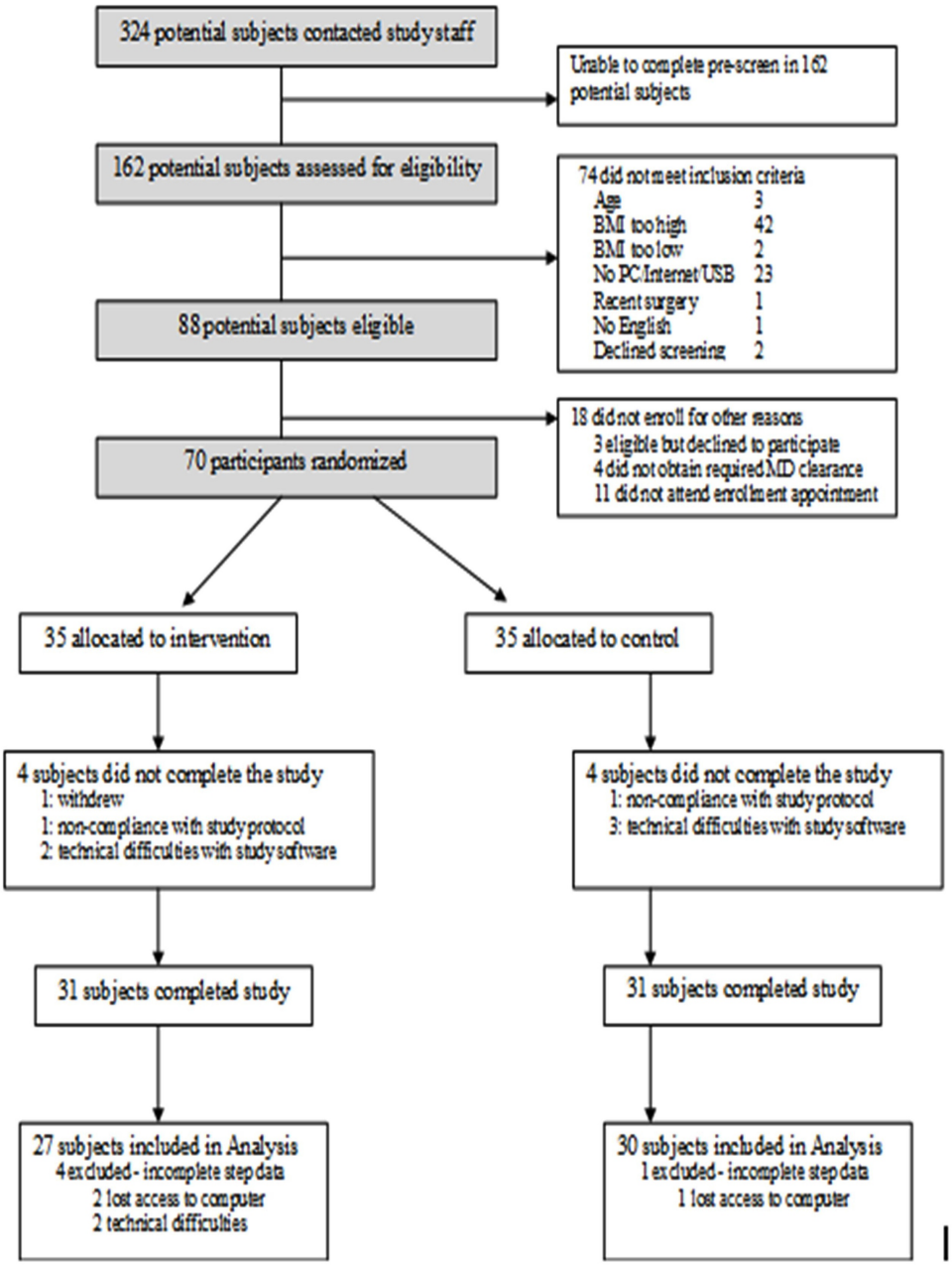
Participants were predominantly female(59/70, 84%), white (53/70, 76%), and college educated (68/70, 97%). Detailed baseline demographic information is reported in Table 1. There were no significant differences in age or baseline BMI between those who completed and those who did not complete the study.

Table 1. Baseline demographics of study participants

	Control (n = 35)	Intervention (n = 35)	*P* value
Female gender, n (%)	28/35 (80)	31/35 (89)	.51
Mean/median age (years)	40.6/41	44.1/9	.11
White race, n (%)	25/35 (71)	28/35 (80)	.57
Mean/median weight (kg)	84.1/82.4	81.3/81.7	.29
Mean/median BMI[a] (kg/m^2)	30.4/29.7	30.2/30.1	.79
College, n (%)	34/35 (97)	34/35 (97)	1.00
Income >US $50,000, n (%)	21/35 (62)	26/35 (74)	.31
Smoker, n (%)	2 (6)	2 (6)	1.0
Survey s (median scores)			
7-day activity (kcal/day)	3071	2890	.13
Self-efficacy	18.9	18.9	.94
Benefits and barriers	136.6	137.1	.88
Social support	35.4	34.6	.62
Technology , n (%)			
Ever used a pedometer	18/35 (51)	14/35 (41)	.47
Use Internet for information on exercise	29/35 (83)	21/35 (62)	.06
Use Internet for information on diet	30/35 (86)	26/35 (77)	.37
Ever used an online fitness program	8/35 (23)	10/35 (29)	.78
Comorbidities , n (%)			
Hypercholesterolemia	6/35 (17)	5/35 (14)	1.0
Hypertension	3/35 (9)	5/35 (14)	.70
Asthma	4/35 (11)	4/35 (11)	1.0

[a] Body mass index.

Figure 2. Flow diagram of participants' progress through the randomized controlled trial. BMI = body mass index; MD = medical doctor.



Outcomes

Figure 3 shows the average step count per period according to group. The average step count in the control group fell significantly from 7174 in P1 to 6149 in P4 (P = .01). In contrast, the intervention participants' mean step count did not change significantly from P1 (6943) to P4 (6943 vs 7024, respectively, P = .85).

Figure 4 charts the percentage change in mean activity levels for each period relative to period 1, only including participants with data for each period (30 control and 27 intervention). The difference seen between groups between P1 and P3 was statistically significant (P = .02), whereas the differences between P1 and P2, and between P1 and P4 did not meet criteria for significance (P = .12 and .07, respectively). A repeated measures analysis of variance, incorporating all data points from this figure, demonstrated that the percentage change in step count across all study periods was significantly different in the intervention versus control arms (P = .02).

We repeated this analysis including those participants who had data missing from one or more of the periods (eg, all completers). The changes from P1 to P2, P3, and P4 were not significant in this analysis (P = .14, .18, and .07 respectively). Due to the shifting sample size in this analysis it was not possible to perform a repeated measures analysis of variance on these data.

There were no statistically significant changes in secondary outcome measures. Mean reduction in BMI was 0.04 and 0.25 in control and intervention participants respectively (P = .44)

Figure 3. Average (SE) daily step count per period. Data shown for all participants who completed the study, including those with missing data.

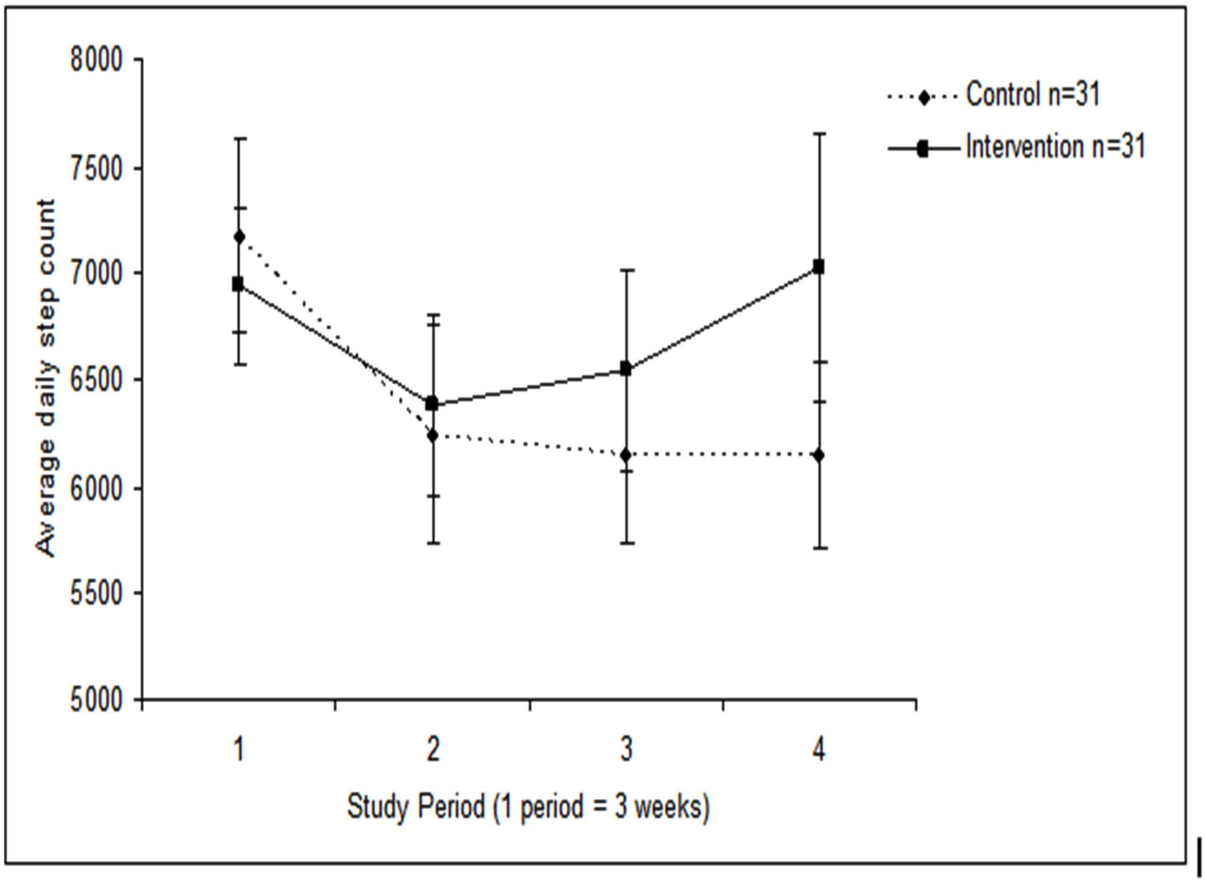
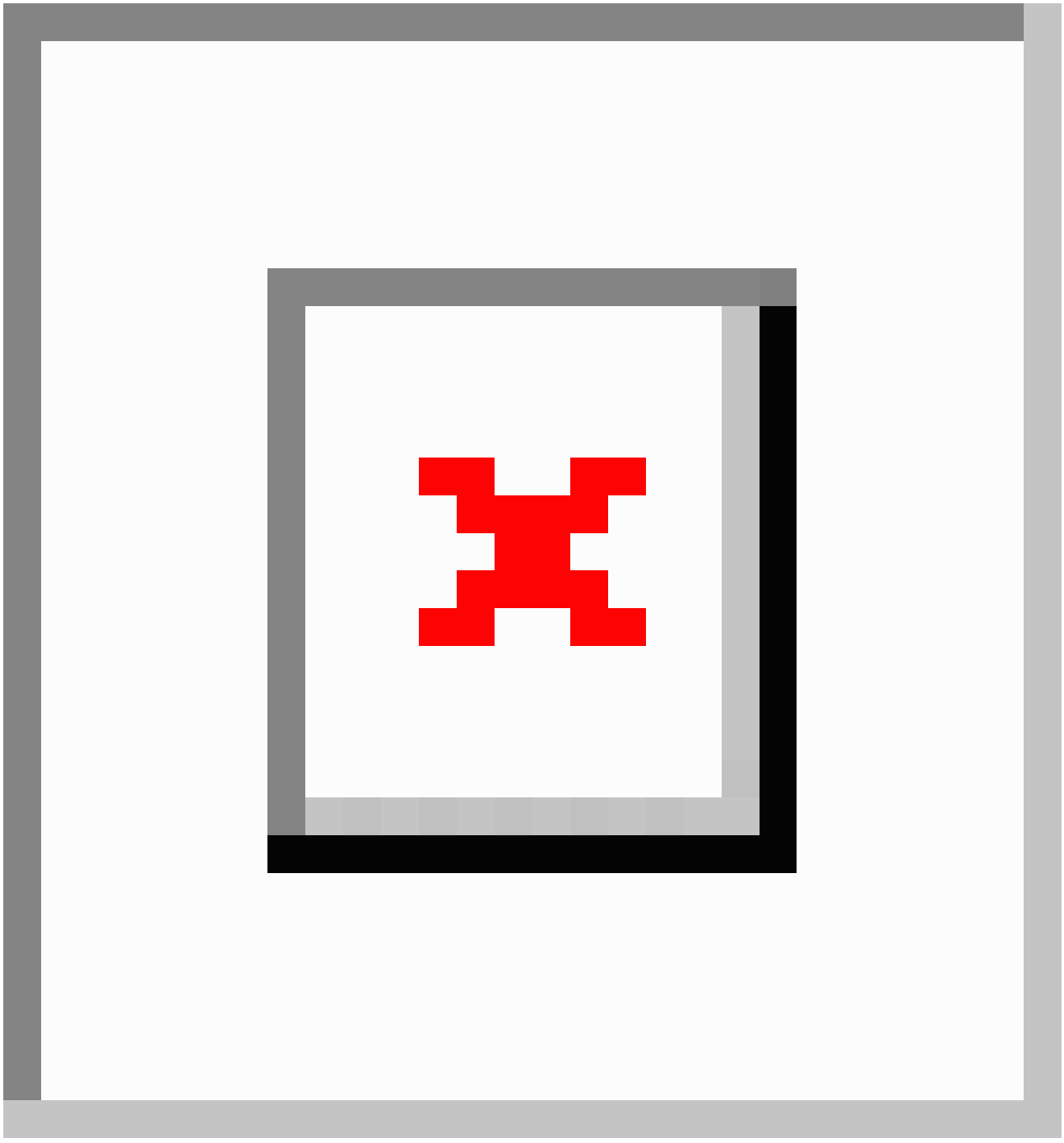


Figure 4. Average (SE) percentage change in step count in each period relative to the first time period (P1) for participants with complete data (ie, no missing periods).



Participants' Adherence to Protocol

The mean number of days that step data were recorded over the course of the study was 73/84 (87%) for the control group and 71/84 (85%) for the intervention group (P = .64). Intervention participants had a mean number of sessions per participant over the course of the study of 28.9 (range 3–63, recommended 36). The mean number of visits per week fell from 2.8 in week 1 to 1.9 in week 12, although this change was not statistically significant (P = .08). There was no significant correlation between the number of sessions intervention participants had with the coach and their performance, in terms of either absolute step increase, absolute step count, or slope of step count change during the intervention.

Participant Benefits and Satisfaction

Both intervention and control participants reported having benefited from taking part in the study (28/30, 93% and 28/31, 90% respectively, P = .67). Self-reported changes by intervention and control participants included exercising more frequently (25/29, 86% vs 21/29, 72%, P = .19) and improved diet and eating habits (13/29, 45% vs 6/29, 21%, P = .05), respectively.

Intervention participants were asked specific questions regarding their interactions with the virtual coach: 18/31 (58%) agreed that the coach motivated them to become more active and 27/31 (87%) reported feeling guilty if they skipped an appointment with the coach.

Discussion

In this trial, providing overweight participants with access to a virtual coach in addition to a pedometer and website appeared to sustain step count over the course of the 12-week study, while step counts in control participants decreased from the start to the end. The percentage change in step count between those in the intervention and those in the control arms, from the start of the study compared with the end of the study, did not reach the threshold for significance ($P = .07$). However, repeated measures analysis found a significant difference when comparing percentage changes in step counts between control and intervention participants over all time points ($P = .02$).

These findings suggest that virtual coaching may be a useful adjunct to existing automated applications designed to promote activity, such as pedometers and online Web programs, and affects participant behavior over and above physiologic monitoring and text- and graph-based feedback. The virtual coach provides an interactive relationship, which is absent in the use of a pedometer or website alone. Previous research has shown that participants can form a social bond with a computer agent, even though they are aware that the agent does not represent an actual human [14]. We observed a somewhat dichotomous response to the virtual coach with some participants reporting negative emotions, such as guilt, anger, or frustration. Some of these emotions may actually promote activity, with the coach functioning as an external conscience or source of accountability, whereas others may erode the effectiveness of the relationship. Although we were not able to demonstrate significant difference in activity based on participants' reactions to the coach, the observed trend suggests that dissatisfaction negatively affects the ability of the coach to promote increased activity. Further investigation into which segments of the patient population are most responsive to this form of automated coaching would allow us to maximize the impact of this intervention.

Figure 4 depicts a decline in activity in controls at the end of the study relative to their activity over the first 3 weeks. It appears that, rather than comparing which interventions increase activity by the largest amounts over extended time periods, interventions should be measured in terms of how well they promote sustained activity over time. Study participants, like the general population, are likely to be most motivated when commencing an activity regimen; maintaining this commitment over time is the challenge. Our study demonstrates that the decline in activity observed in control participants may be overcome through virtual coaching over a 12-week time period. In fact, the mean daily step counts in the last 3 weeks of the study were around 800 steps higher in intervention participants than in control participants, although this difference was not statistically significant. Given that pedometers have been demonstrated to increase step count by around 2000 steps per day [6], this suggests that virtual coaching might offer an additional increase in steps compared with pedometers alone. Longer study would be required to assess the effectiveness of virtual coaching over extended time periods.

Developing effective, automated self-management programs, that offer a relationship and personalized feedback, may prove essential to developing scalable solutions to deal with large populations faced with chronic disease. Technology-based solutions are becoming increasingly feasible as the proportion of Americans with access to cell phones and the Internet rises: current adoption estimates are 85% and 79%, respectively [28]. Internet-based health behavior-change programs may have more potential in the area of weight management than in many other health-related areas [29], as obese people are more likely to participate in follow-up because of a personal and nonstigmatizing approach via the Internet. Internet-based physical activity and weight-management programs have been shown to promote behavior change [30], particularly in occupational health settings [31], but the results haven't always been consistent. A randomized trial of an Internet-based weight-management intervention among military personnel had shown benefits by preventing weight gain [32]. Studies should, however, assess the quality and appropriateness of the tailored advice. For example, a study by Slootmaker et al [33] evaluated the feasibility and effectiveness of a 3-month intervention in which office workers were provided with a personal activity monitor coupled to simple and concise Web-based tailored physical activity advice. They observed no significant intervention effect in the physical activity outcomes at the 8-month follow-up. Whether Internet-based interventions are more effective in helping people sustain their weight loss than in promoting new weight loss is an area of active inquiry [34].

Equitable access to technology is still a further concern, as groups most affected by obesity, such as minorities and low-socioeconomic status groups, are least likely to have access to technology [35]. Even low-cost, text messaging-based interventions delivered on cell phones have shown positive short-term behavioral outcomes [36]. Therefore, providing access to computers at local libraries or creating coaching applications that run on cell phones may be helpful strategies to bridge this divide [37-39].

This application may also be of value in tackling weight and activity issues among the pediatric or adolescent population. These groups are avid users of new technology and may be receptive to automated Web-based programs. To reduce the prevalence of overweight and obesity in these populations, it may also be necessary to decrease time spent in sedentary behaviors, such as leisure-time Internet and computer use [40]. Previous projects in these populations have primarily focused on school-based interventions, although the need to develop and evaluate Internet-based approaches has been flagged as a priority area for research [41].

There were several limitations to this study. Our study participants were primarily white, college-educated women. As a result, it may be difficult to generalize our findings to the wider population of overweight or obese patients who may be less comfortable taking a more active role in managing their health or in using technology. We do not have access to baseline step counts for study participants. We did, however, survey participants about baseline activity levels and found no significant difference. The step counts observed in the first few weeks after enrollment likely reflect an increase from baseline

step counts for both intervention and control participants. It is also difficult to compare the results of this intervention with other pedometer-based programs because, unlike many commercially available pedometers, the ActiPed does not give participants immediate feedback on their current step count.

The length of the study was 12 weeks: ideally benefits of an activity or weight-loss program would be assessed over a longer time period. If this study were to be repeated over a longer time period, the coaching algorithm would need to be expanded to allow for a more variable series of interactions to maintain participants' interest. Some of our participants reported finding the coach repetitive even over a 12-week time period. Conducting the study over a longer period of time, or with more participants, would allow for more robust assessment of highly relevant secondary outcome measures, such as decrease in BMI.

Finally, study staff contacted both intervention and control participants if they were noncompliant with the use of the pedometer or virtual coach, or if they were experiencing technical difficulties. There was more contact with the intervention participants (51 calls) than with control participants (31 calls). This likely reflects the fact that the intervention group had two different technologies to operate. This difference in contact may have had some bearing on the observed effects over the course of the study, but we think this is likely to be minor, as contact was not related to level of step count.

Virtual coaching has many applications beyond promoting activity. Coaching is increasingly recognized as an important component in the management of chronic conditions, such as diabetes and heart disease, and in the promotion of healthy behaviors, such as adherence to medication [42]. Given the growing burden of chronic disease and the shortage of providers [43], such applications may prove useful adjuncts to conventional office-based care. By linking data from home monitoring devices such as scales, blood pressure cuffs, or glucometers to automated coaching programs, health practitioners could encourage patients to develop better self-management skills. Clinician input could be provided if predefined triggers were alerted. Physicians have been slow to adopt information technology tools [44] due to concerns around reimbursement, data overload, and clinical outcomes, so thoughtful integration within existing workflow, including electronic medical record integration, would be necessary to ensure participation.

Broader changes in the health care payment system, such as the shift from visit-based to outcome-based reimbursement, may promote adoption of this type of care-delivery platform by clinicians. Pay-for-performance initiatives promote provider innovation around care delivery. An online coaching platform could be used, either as an adjunct to traditional care or as a stand-alone self-management program for patients.

Conclusion

In this study we demonstrated a sustained level of activity in overweight participants provided with a virtual coach in addition to a pedometer and Web-based feedback, compared with a decline seen in those provided with a pedometer and Web-based program alone. Further work should examine the long-term benefits of virtual coaching and the extension of this application to a wider patient population.

Acknowledgments

We would like to thank Deirdre Neylon, Brian Sweeney, and Laura Pfiefer for their contributions to this project.

This study was funded by a grant from the Information Systems Research Council at Partners Healthcare.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH V1.6 checklist [45].

[PDF File (Adobe PDF File), 811KB - jmir_v14i1e1_app1.pdf]

References

1. Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. Prevalence of overweight and obesity in the United States, 1999-2004. JAMA 2006 Apr 5;295(13):1549-1555. [doi: 10.1001/jama.295.13.1549] [Medline: 16595758]
2. Centers for Disease Control and Prevention. USA.gov. 2009 Dec 7. Overweight and Obesity: Causes and Consequences URL: http://www.cdc.gov/obesity/causes/index.html [accessed 2010-07-27] [WebCite Cache ID 5rXjSpFa6]
3. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). US Department of Health and Human Services, National Institutes of Health. 2010 Feb. Weight-control Information Network (WIN): Statistics Related to Overweight and Obesity URL: http://www.win.niddk.nih.gov/STATISTICS/ [accessed 2010-07-27] [WebCite Cache ID 5rXjOyzrW]
4. Morgan O. Approaches to increase physical activity: reviewing the evidence for exercise-referral schemes. Public Health 2005 May;119(5):361-370. [doi: 10.1016/j.puhe.2004.06.008] [Medline: 15780323]
5. Wing RR, Phelan S. Long-term weight loss maintenance. Am J Clin Nutr 2005 Jul;82(1 Suppl):222S-225S [FREE Full text] [Medline: 16002825]

6. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
7. 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] [Medline: 20584698]
8. Wanner M, Martin-Diener E, Braun-Fahrländer C, Bauer G, Martin BW. Effectiveness of active-online, an individually tailored physical activity intervention, in a real-life setting: randomized controlled trial. J Med Internet Res 2009;11(3):e23 [FREE Full text] [doi: 10.2196/jmir.1179] [Medline: 19666456]
9. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr 29;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
10. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
11. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 7;285(9):1172-1177 [FREE Full text] [Medline: 11231746]
12. Harvey-Berino J, Pintauro S, Buzzell P, Gold EC. Effect of internet support on the long-term maintenance of weight loss. Obes Res 2004 Feb;12(2):320-329. [doi: 10.1038/oby.2004.40] [Medline: 14981225]
13. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an Internet weight loss program. Arch Intern Med 2006;166(15):1620-1625 [FREE Full text] [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
14. Okun BF. Effective Helping: Interviewing and Counseling Techniques. Pacific Grove, CA: Brooks/Cole; 1997.
15. Gelso CJ, Hayes JA. The Psychotherapy Relationship: Theory, Research and Practice. New York, NY: Wiley; 1998.
16. Horvath A, Greenberg, L. Development and validation of the Working Alliance Inventory. J Counsel Psychol 1989;36(2):223-233. [doi: 10.1037/0022]
17. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res 2010;12(4):e71 [FREE Full text] [doi: 10.2196/jmir.1338] [Medline: 21169160]
18. Gabriele JM, Carpenter BD, Tate DF, Fisher EB. Directive and nondirective e-coach support for weight loss in overweight adults. Ann Behav Med 2011 Apr;41(2):252-263. [doi: 10.1007/s12160-010-9240-2] [Medline: 21108032]
19. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
20. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
21. Moy ML, Garshick E, Matthess KR, Lew R, Reilly JJ. Accuracy of uniaxial accelerometer in chronic obstructive pulmonary disease. J Rehabil Res Dev 2008;45(4):611-617 [FREE Full text] [Medline: 18712646]
22. Jonas S. Regular Exercise: A Handbook for Clinical Practice. New York, NY: Springer; 1995.
23. King AC. The coming of age of behavioral research in physical activity. Ann Behav Med 2001;23(4):227-228. [Medline: 11761339]
24. Washburn RA, Jacobsen DJ, Sonko BJ, Hill JO, Donnelly JE. The validity of the Stanford Seven-Day Physical Activity Recall in young adults. Med Sci Sports Exerc 2003 Aug;35(8):1374-1380. [doi: 10.1249/01.MSS.0000079081.08476.EA] [Medline: 12900693]
25. Prochaska JO, Redding CA, Harlow LL, Rossi JS, Velicer WF. The transtheoretical model of change and HIV prevention: a review. Health Educ Q 1994;21(4):471-486. [Medline: 7843978]
26. Resnick B, Jenkins LS. Testing the reliability and validity of the Self-Efficacy for Exercise scale. Nurs Res 2000;49(3):154-159. [Medline: 10882320]
27. Sechrist KR, Walker SN, Pender NJ. Development and psychometric evaluation of the exercise benefits/barriers scale. Res Nurs Health 1987 Dec;10(6):357-365. [Medline: 3423307]
28. Horrigan JB. Home broadband adoption 2009. Washington, DC: Pew Internet & American Life Project; 2009 Jun 17. URL: http://pewresearch.org/pubs/1254/home-broadband-adoption-2009 [accessed 2010-07-27] [WebCite Cache ID 5rXkSsd6U]
29. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
30. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
31. Ware LJ, Hurling R, Bataveljic O, Fairley BW, Hurst TL, Murray P, et al. Rates and determinants of uptake and use of an internet physical activity and weight management program in office and manufacturing work sites in England: cohort study. J Med Internet Res 2008;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]

32. Hunter CM, Peterson AL, Alvarez LM, Poston WC, Brundige AR, Haddock CK, et al. Weight management using the internet a randomized controlled trial. Am J Prev Med 2008 Feb;34(2):119-126. [doi: 10.1016/j.amepre.2007.09.026] [Medline: 18201641]
33. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, Van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
34. Stevens VJ, Funk KL, Brantley PJ, Erlinger TP, Myers VH, Champagne CM, et al. Design and implementation of an interactive website to support long-term maintenance of weight loss. J Med Internet Res 2008;10(1):e1 [FREE Full text] [doi: 10.2196/jmir.931] [Medline: 18244892]
35. Kontos EZ, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. J Med Internet Res 2007;9(4):e31 [FREE Full text] [doi: 10.2196/jmir.9.4.e31] [Medline: 17951215]
36. 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]
37. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
38. 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] [Medline: 19141433]
39. Lefebvre C. Integrating cell phones and mobile technologies into public health practice: a social marketing perspective. Health Promot Pract 2009 Oct;10(4):490-494. [doi: 10.1177/1524839909342849] [Medline: 19809002]
40. Vandelanotte C, Sugiyama T, Gardiner P, Owen N. Associations of leisure-time internet and computer use with overweight and obesity, physical activity and sedentary behaviors: cross-sectional study. J Med Internet Res 2009;11(3):e28 [FREE Full text] [doi: 10.2196/jmir.1084] [Medline: 19666455]
41. Salmon J, Booth ML, Phongsavan P, Murphy N, Timperio A. Promoting physical activity participation among children and adolescents. Epidemiol Rev 2007;29:144-159. [doi: 10.1093/epirev/mxm010] [Medline: 17556765]
42. Jelinek M, Vale MJ, Liew D, Grigg L, Dart A, Hare DL, et al. The COACH program produces sustained improvements in cardiovascular risk factors and adherence to recommended medications-two years follow-up. Heart Lung Circ 2009 Dec;18(6):388-392. [doi: 10.1016/j.hlc.2009.06.001] [Medline: 19648058]
43. Bowman RC. Measuring primary care: the standard primary care year. Rural Remote Health 2008;8(3):1009 [FREE Full text] [Medline: 18785798]
44. Grant RW, Campbell EG, Gruen RL, Ferris TG, Blumenthal D. Prevalence of basic information technology use by U.S. physicians. J Gen Intern Med 2006 Nov;21(11):1150-1155. [doi: 10.1111/j.1525-1497.2006.00571.x] [Medline: 16879417]
45. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

BMI: body mass index
PAR-Q: Physical Activity Readiness Questionnaire

Edited by G Eysenbach; submitted 10.08.10; peer-reviewed by R Hurling, C Bexelius, F Slinde; comments to author 21.09.10; revised version received 27.10.11; accepted 07.11.11; published 26.01.12.

Please cite as:
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
URL: http://www.jmir.org/2012/1/e1/
doi:10.2196/jmir.1629
PMID:22281837

Original Paper

Is a Severe Clinical Profile an Effect Modifier in a Web-Based Depression Treatment for Adults With Type 1 or Type 2 Diabetes? Secondary Analyses From a Randomized Controlled Trial

Kim MP van Bastelaar[1,2], PhD; François Pouwer[3], PhD; Pim Cuijpers[4,5], PhD Prof; Heleen Riper[4,5,6,7], PhD; Jos WR Twisk[5,8,9], PhD Prof; Frank J Snoek[1,5], PhD Prof

[1]Department of Medical Psychology, VU University Medical Centre, Amsterdam, Netherlands

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

[3]Centre of Research on Psychology in Somatic Diseases (CoRPS), Department of Medical Psychology & Neuropsychology, University of Tilburg, Tilburg, Netherlands

[4]Department of Clinical Psychology, VU University, Amsterdam, Netherlands

[5]EMGO+ Institute for Health and Care Research, VU University, Amsterdam, Netherlands

[6]GGZinGeest, Amsterdam, Netherlands

[7]Institute of Psychology, Leuphana University, Lüneburg, Germany

[8]Department of Epidemiology and Biostatistics, VU University Medical Centre, Amsterdam, Netherlands

[9]Faculty of Health Sciences, VU University, Amsterdam, Netherlands

Corresponding Author:
Frank J Snoek, PhD Prof
Department of Medical Psychology
VU University Medical Centre
Van der Boechorststraat 7
Amsterdam, 1081 BT
Netherlands
Phone: 31 204448228
Fax: 31 204448230
Email: fj.snoek@vumc.nl

Abstract

Background: Depression and diabetes are two highly prevalent and co-occurring health problems. Web-based, diabetes-specific cognitive behavioral therapy (CBT) depression treatment is effective in diabetes patients, and has the potential to be cost effective and to have large reach. A remaining question is whether the effectiveness differs between patients with seriously impaired mental health and patients with less severe mental health problems.

Objective: To test whether the effectiveness of an eight-lesson Web-based, diabetes-specific CBT for depression, with minimal therapist support, differs in patients with or without diagnosed major depressive disorder (MDD), diagnosed anxiety disorder, or elevated diabetes-specific emotional distress (DM-distress).

Methods: We used data of 255 patients with diabetes with elevated depression scores, who were recruited via an open access website for participation in a randomized controlled trial, conducted in 2008–2009, comparing a diabetes-specific, Web-based, therapist-supported CBT with a 12-week waiting-list control group. We performed secondary analyses on these data to study whether MDD or anxiety disorder (measured using a telephone-administered diagnostic interview) and elevated DM-distress (online self-reported) are effect modifiers in the treatment of depressive symptoms (online self-reported) with Web-based diabetes-specific CBT.

Results: MDD, anxiety disorder, and elevated DM-distress were not significant effect modifiers in the treatment of self-assessed depressive symptoms with Web-based diabetes-specific CBT.

Conclusions: This Web-based diabetes-specific CBT depression treatment is suitable for use in patients with severe mental health problems and those with a less severe clinical profile.

ClinicalTrial: International Standard Randomized Controlled Trial Number (ISRCTN): 24874457; http://www.controlled-trials.com/ISRCTN24874457 (Archived by WebCite at http://www.webcitation.org/63hwdviYr)

(J Med Internet Res 2012;14(1):e2) doi:10.2196/jmir.1657

KEYWORDS

Diabetes mellitus, type 1; diabetes mellitus, type 2; depression; behavior therapy; cognitive therapy; depressive disorder; adults; psychology

Introduction

With an estimated world prevalence of 285 million people, diabetes mellitus has reached epidemic levels globally [1]. Affecting 10% to 20% of the adult diabetes patients, depression is to be regarded as a common comorbid health problem that negatively affects quality of life and diabetes outcomes, and increases mortality [2,3]. Treating depression in diabetes is therefore of great importance.

A recent meta-analysis has shown that depression in diabetes patients can be effectively treated with various antidepressant treatments, showing the highest effect sizes for psychological treatment (Cohen *d* = –0.58, 95% CI, –0.77 to –0.39), compared with pharmacologic treatment (Cohen *d* = –0.47, 95% CI, –0.66 to –0.27) or collaborative care (Cohen *d* = –0.29, 95% CI, –0.43 to –0.16) [4].

Nevertheless, in a substantial portion of diabetes patients, comorbid depression remains untreated. Underrepresentation of complaints, underrecognition of depressive symptoms by health care providers, and inadequate referral can account for this [5,6]. Another reason for untreated depression in diabetes patients is the negative stigma of mental health care among patients who are treated in physical health care, or that they do not feel at home in a mental health care setting [7,8]. It has also been suggested that, since generic and disease-specific emotional distress are not the same, we need to tailor interventions to the specific needs of this subgroup of patients [9,10]. A depression treatment that specifically addresses elevated diabetes-specific emotional distress (DM-distress) could help overcome this last barrier to treatment. The VU University Medical Center in Amsterdam, The Netherlands, has recently developed such a cognitive behavioral therapy (CBT) depression intervention, specifically tailored to the needs of diabetes patients by incorporating diabetes-specific topics, such as coping strategies for diabetes-specific issues [11].

Providing psychological interventions via the Internet could help overcome barriers to treatment related to travelling distance and time—for example, it has the potential to avoid reluctance to seek therapy among patients who are ashamed of needing psychological help, and it allows patients to work at home, in their own pace and familiar environment, while saving them time, and the burden and cost of travelling. Therefore, an Internet-administered intervention has the potential to have a broad reach. Internet-administered therapy can also save therapists time, thus reducing waiting lists [12]. Providing psychological interventions via the Internet can be a major advantage specifically for diabetes patients, since they already spend much time in (somatic) health care, and severe diabetes complications can cause physical impairments, causing difficulties in patients' mobility. Considering these advantages, the diabetes-specific depression intervention, called Diabetergestemd.nl (DbG.nl), was offered via the Internet and tested in a randomized controlled trial (RCT) [13]. The DbG.nl intervention was found to be significantly more effective than a waiting-list control condition [14].

A commonly heard criticism in studies regarding Web-based CBT depression treatment is that most studies do not differentiate between elevated symptoms of depression (subclinical depression) and diagnosed depression in the strict sense, or major depressive disorder (MDD) [15]. This causes clinicians to need more convincing evidence that online CBT can help their patients with subclinical depression, but also those with MDD, especially since the current guidelines for depression treatment indicate that patients with subclinical depression warrant a different treatment (low-intensity psychosocial interventions, such as guided self-help) from patients with MDD (high-intensity psychosocial intervention, such as individual CBT) [16]. In the RCT studying the effectiveness of DbG.nl, a diagnostic instrument was administered, which enabled us to make a clear distinction between patients with subclinical depression and patients with MDD. Examination of potential differences in effectiveness between both subgroups provides important information regarding the potential utility of the intervention from a public health perspective.

The elevated prevalence of anxiety disorders in type 1 and type 2 diabetes in comparison with prevalence rates in the general population has been demonstrated in a systematic review [17]. Since studies on the prevalence of co-occurring anxiety and mood disorders in diabetes patients have yielded mixed results [18], we were interested in exploring the co-occurrence of anxiety disorders in patients with MDD. Furthermore, the treatment literature indicates that the combination of MDD and anxiety disorder is more difficult to treat than MDD alone [19]. Therefore, we aimed to examine the effect modification of anxiety disorders in our study sample.

An important issue to consider in the context of psychological interventions for people with diabetes is the role of DM-distress and the need to accurately differentiate DM-distress from general emotional distress [20]. Previous reports have emphasized that diabetes patients with elevated symptoms of depression are not all necessarily clinically depressed, but rather may have high levels of diabetes-related distress [21,22]. Our RCT allowed us to test whether the online diabetes-specific depression intervention was more or less effective in patients with baseline elevated DM-distress than in those with lower levels of DM-distress.

To summarize, we set out to answer the following questions: does the effectiveness of a Web-based, diabetes-specific CBT depression intervention differ (1) for patients with or without MDD, (2) for patients with or without an anxiety disorder, and (3) for patients with or without elevated DM-distress? We hypothesized that Web-based, diabetes-specific CBT is more

effective in diabetes patients with MDD, anxiety disorder, or elevated DM-distress.

To the best of our knowledge, this is the first study that used data from an RCT to perform secondary analyses to test effect modification in subgroups of diabetes patients regarding the effectiveness of a Web-based, diabetes-specific CBT depression treatment.

Methods

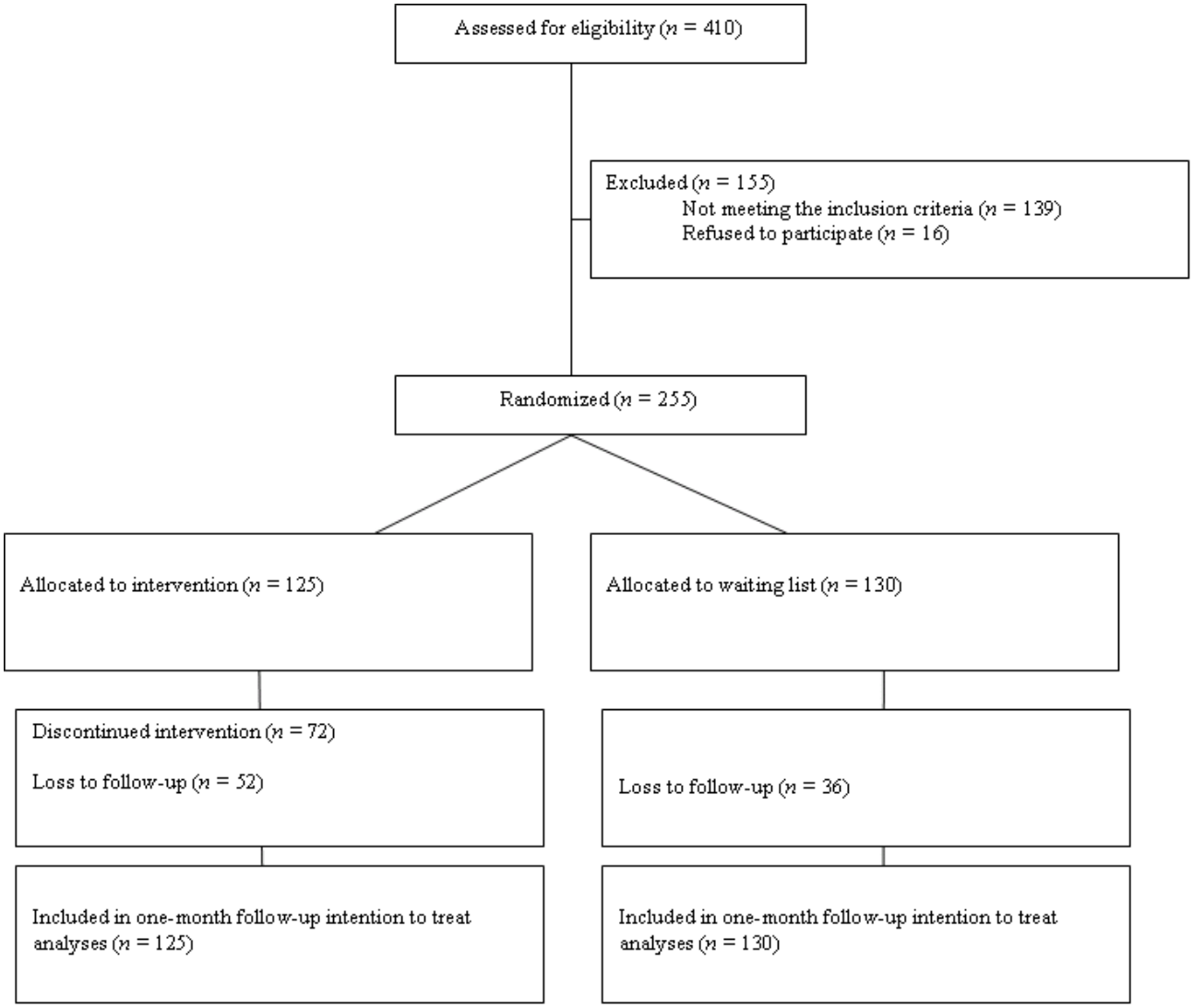
Participants and Procedure

Patients for an RCT were recruited from July 2008 trough September 2009. We randomly assigned 255 adult diabetes patients with elevated depressive symptoms (having a score of 16 or higher on the Center for Epidemiologic Studies Depression scale [CES-D]) to the Web-based diabetes-specific depression intervention (n = 125) or a 12-week waiting-list control group (n = 130) (see Figure 1) [13]. Exclusion criteria were a history of suicide attempt(s) or current suicidal ideation; bipolar depression or psychotic disorder; pregnancy; recent loss of a significant other (<6 months ago); and insufficient Internet literacy. Patients were recruited via advertisements in various media and could sign up for participation in the study via an open access study website. The study was advertised as a study performed by the VU University Medical Center for testing the effectiveness of diabetes-specific online depression treatment. It was mentioned that attending was cost-free and no financial reward was provided.

Patients could individually sign up for participation in the study through an open access study website. Written informed consent was obtained by mail and included information on the study, permission for anonymous data use, and permission to contact the patient's general practitioner and treating diabetes physician for obtaining data on diabetes. After having signed the informed consent, patients were invited to fill out the baseline assessment through a personal online questionnaire, and they received a telephone-administered diagnostic interview.

Individual randomization by computer was used to assign participants to either the experimental or control condition, at an individual level, using a 1:1 allocation ratio. Due to the nature of the study (waiting-list controlled) it was not possible to blind patients to treatment allocation. The sample size was calculated based on the expected difference in the primary outcome variable (ie, depressive symptoms). Based on a statistical power of 80%, with an alpha of .05, we required 100 participants in each group to be able to detect differences with an effect size of 0.35. The design of the RCT on the effectiveness of the Web-based diabetes-specific depression intervention has been described in more detail elsewhere [13]. Data of the RCT were used to perform secondary analyses on effect modification. All randomly assigned participants were analyzed. The Medical Ethics Committee of the VU University Medical Center approved the study. The results of the RCT have been described elsewhere [14]. In short, the intervention was effective in reducing depressive symptoms and diabetes-specific emotional distress.

Figure 1. CONSORT study flowchart.



Intervention

Participants assigned to the intervention group individually attended the online course, based on the principles of CBT. The intervention DbG.nl (www.diabetergestemd.nl) was developed by the VU University Medical Center in collaboration with the Trimbos Institute. DbG.nl was based on the effective Web-based CBT depression intervention Color Your Life [22], the Internet version of Lewinsohn's effective and well-known Coping with Depression course [23]. DbG.nl follows the same format as Color Your Life, putting emphasis on the following skills: relaxation, cognitive restructuring (including worrying), positive reinforcement, assertiveness, communication skills, and increasing the number of pleasant activities. In short, the course consists of eight consecutive weekly lessons, consisting of psychoeducation and focused on skills such as relaxation, cognitive restructuring (including worrying), positive reinforcement, social skills, and increasing the number of activities that are pleasant to the patient. The course contained written and spoken information and homework assignments (see Figure 2 for a screenshot) with one-time incorporated email feedback for each lesson from a coach (qualified psychologist). Patients were advised to go through one lesson per week. In case we did not receive their homework, patients were sent reminders after 1 week and after 2 weeks. If we received no reply within 3 weeks, we sent participants an email stating that we had to assume that they were no longer interested in the intervention, and we invited them to fill out the postmeasurement. However, if they were still interested, they were invited to reenter the course. After the course, patients were invited to fill out feedback forms, and we interviewed several patients by telephone.

Patients attended nonanonymously, so having multiple identities was not possible. Coaches only knew patients' names, and all personal data were omitted after the patients' medical data were obtained from their treating physician (to which the patients consented) and the participant completed or withdrew from the course. Coaches were blinded to treatment allocation of participants.

Figure 2. Screenshot of a lesson in the web-based diabetes-specific course.

Sample Characteristics

Baseline characteristics of our study sample were self-reported as part of the online assessment. Previous analyses confirmed successful randomization: intervention and control groups did not statistically significantly differ regarding demographic or clinical characteristics at baseline (all $P \geq .05$) [14]. Of those randomly assigned to the intervention group, 53 (42%) completed the entire eight-lesson course, 30 (24%) completed no lesson at all, and 7 (6%) never logged into the course. Other participants dropped out equally divided during the course.

Outcome Measure

Depressive symptoms were self-assessed online with the Dutch version of the CES-D. The CES-D is a validated self-report screening instrument that measures the frequency with which participants have experienced specific depressive symptoms within the preceding week. The questionnaire contains 20 items assessed on a 4-point Likert scale. The total score can range from 0 to 60, where higher scores indicate more depressive symptoms. In Dutch samples, a cut-off point of 16 or higher is generally accepted to indicate clinical depression [24].

Potential Effect Modifiers

Based on our research questions, we selected the following three potential effect modifiers: MDD (yes/no), anxiety disorder (yes/no), and high versus low level of DM-distress.

To diagnose MDD and anxiety disorder, the Dutch, we administered by telephone the computerized version of the World Health Organization Composite International Diagnostic Interview (WHO CIDI-auto), a fully structured psychiatric diagnostic interview that assesses diagnostic criteria of mental disorders according to the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition. The WHO CIDI-auto is a computerized version of the WHO CIDI, a qualified substitution of the face-to-face interview [25]. Since questions and routes are fully specified, no clinical judgment is required. Interviewers were masters students in clinical psychology of the VU University in Amsterdam, trained in the administration of the WHO CIDI-auto by telephone.

DM-distress was measured using the Dutch-validated Problem Areas in Diabetes scale (PAID), which was self-assessed online [26]. We compared the effectiveness of our intervention in patients with elevated DM-distress (using the cut-off of PAID ≥40) with those without elevated DM-distress (PAID <40).

Statistical Analyses

To test effect modification, the course of depressive symptoms at baseline, posttreatment, and 1-month follow-up between the intervention and control groups was compared for each potential effect modifier. We performed generalized estimating equation (GEE) analyses using 3-way interaction terms (group × time × potential effect modifier) to examine whether having an MDD

(yes/no) diagnosis, anxiety disorder (yes/no) diagnosis, and elevated DM-distress (PAID ≥40/<40) were significant effect modifiers in the treatment effect. All analyses were corrected for baseline depression scores to gain insight into the relative degree of change, and for time between baseline measurement and postmeasurement. Also, all potential effect modifiers were examined on baseline differences for all of the measured sociodemographic and clinical variables, and all analyses were corrected for these differences.

In the RCT data, overall study attrition was 32% at postassessment and 35% at 1-month follow-up. Since study attrition was higher in noncompleters of the course at the 1-month follow-up than in completers of the course (63% vs 13%; $P < .001$), attrition was not completely at random. We therefore imputed missing data using the state-of-the-art multiple imputation technique with Stata 10.0 software (StataCorp LP, College Station, TX, USA). Multiple imputation minimally alters variance of data and thus provides best estimates of true data [27]. All further statistical analyses were performed using complete data, with either Stata 10.0 or SPSS 15.0 software (IBM Corporation, Somers, NY, USA). All results were based on intention-to-treat analyses.

Results

Baseline Characteristics

Baseline sociodemographic and clinical characteristics of our study sample are presented in Table 1, and have been described in more detail elsewhere [14].

Table 1. Baseline sociodemographic and clinical characteristics of study sample at baseline

Characteristics	All patients (n = 255)	CBT[a] participants (n = 125)	Waiting-list control participants (n = 130)	*P* value
Sociodemographics				
Age (years), mean (SD)	50 (12)	48 (12)	51 (12)	.51
Women, n (%)	155 (60.7)	82 (66)	73 (56)	.12
White, n (%)	227 (89.0)	110 (88)	117 (90)	.87
Marital state: with partner, n (%)	199 (78.0)	99 (79)	100 (77)	.66
Education level, n (%)				
No formal qualifications	8 (3)	5 (5)	3 (3)	.44
High school or lower/middle vocational qualifications	136 (53.3)	70 (56)	66 (51)	.40
College qualifications or more	111 (43.6)	50 (40)	61 (47)	.27
Clinical characteristics				
Depressive symptoms (CES-D[b], range 16–60), mean (SD)	28 (7)	29 (7)	28 (7)	.50
Diabetes-specific emotional distress (PAID[c], range 0–100), mean (SD)	40 (19)	42 (19)	38 (19)	.05
Type 2 diabetes, n (%)	141 (55)	66 (53)	75 (58)	.43
Mean HbA_{1c} level[d], %	7.4 (1.3)	7.4 (1.6)	7.3 (1.6)	.36
Self-reported diabetes complications, n (%)				
Neuropathy	25 (10)	11 (9)	14 (11)	.23
Nephropathy	11 (4)	5 (4)	6 (5)	.69
Retinopathy	30 (12)	17 (14)	13 (10)	.69
Foot ulcer	21 (8)	9 (7)	12 (9)	.24
Diagnosis of depressive disorder (WHO CIDI-auto)[e], n (%)				
MDD[f]	146 (57)	71 (57)	75 (58)	.89
MDD, single episode, mild	62 (24)	29 (23)	33 (25)	.68
MDD, single episode, moderate	48 (19)	26 (21)	22 (17)	.43
MDD, single episode, severe	21 (8)	9 (7)	12 (9)	.56
MDD, recurrent episode, mild	8 (3)	4 (3)	4 (3)	1.00
MDD, recurrent episode, moderate	4 (2)	1 (1)	3 (2)	.33
MDD, recurrent episode, severe	3 (1)	2 (2)	1 (1)	.54
Dysthymic disorder	28 (11)	13 (10)	15 (12)	.68
Diagnosis of anxiety disorder, n (%)	95 (37)	43 (34)	52 (40)	.36
Generalized anxiety disorder	59 (23)	25 (20)	34 (26)	.24
Social phobia	24 (9)	12 (9)	12 (9)	.92
Panic disorder	12 (5)	4 (3)	8 (6)	.27
Panic disorder with agoraphobia	5 (2)	3 (2)	2 (2)	.62
Agoraphobia	11 (4)	3 (2)	8 (6)	.14
Specific phobia	42 (16)	20 (16)	22 (17)	.83
Blood-injection-injury type	22 (9)	15 (12)	7 (5)	.06
Environment type	15 (6)	6 (5)	9 (7)	.47

Characteristics	All patients (n = 255)	CBT[a] participants (n = 125)	Waiting-list control participants (n = 130)	*P* value
Situational type	9 (4)	5 (4)	4 (3)	.69
Animal type	6 (3)	1 (1)	5 (4)	.12

[a] Cognitive behavioral therapy.

[b] Center for Epidemiologic Studies Depression scale.

[c] Problem Areas In Diabetes scale.

[d] Glycosylated hemoglobin.

[e] World Health Organization Composite International Diagnostic Interview.

[f] Major depressive disorder.

Potential Effect Modifiers

As shown in Table 1, over half of the patients in our study sample (n = 146, 57.3%) had an MDD diagnosis, of whom the majority (131/146, 89.7%, had a single episode of MDD, not a recurrent depression. About half of the patients with an MDD comorbidly had an anxiety disorder (69/146, 47%) and about half had comorbid elevated DM-distress (80/146, 55%) (Table 2). Furthermore, of patients with an MDD diagnosis, a higher percentage had type 2 diabetes (94/164, 64% vs 47/109, 43%, $P < .001$) and a lower percentage used antidepressant medication (7/109, 6% vs 21/146, 14%, $P < .001$). MDD patients reported higher baseline symptoms of depression, with a mean (SD) CES-D of 30 (7) versus 26 (7), $P < .001$, and higher levels of DM-distress, mean (SD) PAID 42 (20) versus 37 (17), $P = .02$. About a third of the total study sample (n = 95, 37%) had an anxiety disorder diagnosed. Patients with an anxiety disorder diagnosis had higher baseline depressive symptoms, with a mean (SD) CES-D of 31 (8) versus 27 (7), $P < .001$, and DM-distress, mean (SD) PAID 48 (18) versus 36 (18), $P < .001$, and did not differ significantly on any sociodemographic variable.

Table 2. Prevalence (%) of diagnosed depression, diagnosed anxiety disorder by elevated and low diabetes-specific emotional distress (40 < PAID[a] ≥ 40) among the study population (n = 255)

	Study population (n = 255)	Elevated diabetes-specific emotional distress	Low diabetes-specific emotional distress
MDD [b]	146 (57)		
Anxiety[c]	69 (27)	46 (18)	23 (9)
No anxiety	77 (30)	34 (13)	43 (17)
No MDD	109 (43)		
Anxiety	26 (10)	17 (7)	9 (4)
No anxiety	83 (33)	30 (12)	53 (21)
Total		127 (50)	128 (50)

[a] Problem Areas In Diabetes scale.

[b] Major depressive disorder measured with the computerized version of the World Health Organization Composite International Diagnostic Interview (WHO CIDI-auto).

[c] Anxiety disorder (WHO CIDI-auto).

Half of our study sample (n = 127, 49.8%) had elevated DM-distress. The patients with elevated DM-distress were younger, with a mean (SD) age of 47 (3) versus 53 (12), $P < .001$; were likelier to be female (85/127, 67% vs 70/128, 55%, $P = .045$); had higher baseline depression scores of mean (SD) CES-D 31 (7) vs 26 (6), $P < .001$; and were more likely to have an anxiety disorder (63/127, 50% vs 32/128, 25%, $P < .001$) than those without elevated DM-distress.

Potential Effect Modifiers of the Treatment Effect

GEE analysis showed that having a diagnosis of MDD ($P = .49$) was not a significant effect modifier in the treatment effect on depressive symptoms (Table 3). In other words, we did not find significant differences in reduction of depressive symptoms for the intervention group versus control group, for patients with MDD compared with patients without MDD. This is after correcting for baseline differences in type of diabetes, use of antidepressant medication, depressive symptoms, and DM-distress, and the time between pre- and posttreatment.

Table 3. Intention-to-treat analyses (n = 125/130) of effectiveness of a Web-based diabetes-specific depression therapy on symptoms of depression as assessed by a CES-D[a] score, testing effect modification by depression status, anxiety disorder, or high level of diabetes-specific emotional distress, in a cognitive behavioral therapy (CBT) intervention versus waiting-list (WL) control group[b]

	Pretreatment		Post treatment		1-month follow-up		*P* value
	CBT	WL	CBT	WL	CBT	WL	
MDD[c,d]	30 (7)	30 (7)	21 (11)	24 (9)	20 (12)	24 (10)	.49
No MDD	27 (7)	26 (7)	18 (9)	21 (8)	19 (10)	20 (8)	
Anxiety disorder[d]	32 (7)	31 (8)	23 (11)	25 (9)	22 (11)	25 (10)	.71
No anxiety disorder	27 (7)	26 (6)	19 (10)	21 (8)	19 (11)	21 (8)	
Elevated diabetes-specific emotional distress (PAID ≥40)[e]	31 (7)	31 (8)	22 (11)	24 (9)	21 (12)	24 (9)	.92
No elevated diabetes-specific emotional distress (PAID >40)[e]	26 (7)	26 (6)	18 (10)	22 (9)	18 (10)	21 (9)	

[a] Center for Epidemiologic Studies Depression scale.

[b] Data are given as mean (SD). Statistical tests relied on generalized estimating equation analyses. *P* values indicate level of significance of effect modification. All analyses are adjusted for baseline CES-D scores, baseline between-group differences on sociodemographic variables, and differences in time between pretreatment and posttreatment. Data are uncorrected.

[c] Major depressive disorder.

[d] Diagnosed using the computerized version of the World Health Organization Composite International Diagnostic Interview (WHO CIDI-auto).

[e] Problem Areas In Diabetes scale.

Similarly, GEE analysis revealed that having an anxiety disorder diagnosis was not a significant effect modifier (*P* = .71) (Table 3). This is after correcting for baseline differences in depressive symptoms and DM-distress, and the time between pre- and posttreatment.

Also, having elevated DM-distress (*P* = .92), was not a significant effect modifier in the treatment effect on depressive symptoms (Table 3), after correcting for age, gender, baseline depression scores, baseline diagnosis anxiety disorder, and the time between pre- and posttreatment.

Discussion

In this study, we aimed to answer the following questions: does the effectiveness of a Web-based diabetes-specific CBT depression intervention differ (1) for patients with or without MDD, (2) for patients with or without an anxiety disorder, and (3) for patients with or without elevated DM-distress? Secondary analyses from an RCT comparing Web-based diabetes-specific depression treatment versus a waiting-list control in patients with type 1 and type 2 diabetes with comorbid depression show that a diagnosis of MDD or anxiety disorder, or reporting high DM-distress does not significantly modify the effect of Web-based diabetes-specific CBT depression treatment. These findings thus suggest that there is no reason to exclude patients with more severe depression or anxiety from participating in what is often considered a first line of treatment for patients with mild depression, following a stepped-care approach [28]. When referring patients to this Web-based, diabetes-specific depression treatment, screening for depressive symptoms using a questionnaire seems advisable, omitting the need to strictly diagnose MDD. It is at this stage unclear for whom the intervention is contraindicated. In our study, patients were excluded in case of psychotic features, suicidal ideation or a history of suicide attempts, or previous admission to a psychiatric hospital for depression treatment. It seems advisable, at minimum, to check for suicidal ideation and psychotic features as exclusion criteria.

Interestingly, roughly half of the participating diabetes patients with comorbid depression reported high disease-specific distress, which was found not to be an effect modifier. This suggests that DbG.nl is suitable for patients with and without high diabetes-related distress. In a future study, comparing the diabetes-specific intervention with a generic Web-based depression intervention on effectiveness and attractiveness from the patient's perspective would be of great importance.

In interpreting the results of our study, we should acknowledge several strengths and limitations of the study. The most important strengths are the design of the study, being an RCT, and the innovative character of the study. This is the first study that tested the effectiveness of Web-based, diabetes-specific depression treatment in different subgroups of patients. Moreover, we administered a diagnostic interview and used validated instruments for measuring depressive symptoms and DM-distress.

An important limitation that needs mentioning is that examining the effect modification was not the primary aim of our RCT. Therefore, this study was not powered to detect significant differences in effect between patients with and without MDD, anxiety disorder, and elevated DM-distress. Yet, due to the relatively large sample size of the study, we had substantial subgroups of patients to compare (n = 146 patients with MDD, n = 109 patients with anxiety disorder, and n = 127 patients with elevated DM-distress). We were unable to test whether specifically MDD was more difficult to treat when patients had comorbid anxiety, but instead examined the effect modification in the full sample of subclinically depressed and MDD patients.

Since in our study sample a substantial group had both MDD and anxiety disorder, this stresses the importance of testing the effect modification of anxiety disorder in a sample of diabetes patients with MDD.

Regarding external validity of our results, we should take into account that our sample consisted largely of white, relatively well-controlled diabetes patients, including only a few less-educated people (about 8/255, 3%) and patients older than 65 years (24/255, 9%). Considering the increased prevalence of diabetes in ethnic populations, older people, and those with lower social economic status, further research is warranted to test the effectiveness of our program in more diverse populations. We also observed that only a few of the participants with MDD had recurrent depression, even though MDD has been shown to be more recurrent in diabetes patients [29]. Perhaps the underrepresentation of patients with recurrent depressive episodes in our study can be explained by these patients already having tried several forms of depression treatment and, and therefore being less willing to try a new form of therapy (Web-based therapy). Future studies should make an effort to attract diabetes patients with recurrent depression in order to test the effectiveness of Web-based therapy.

Conclusions

Findings from this study provide the first evidence to suggest that Web-based diabetes-specific depression treatment is effective in patients with mild to more severe depression, with or without comorbid anxiety disorder. Information on differential effects of Web-based therapy is vital to make evidence-based recommendations regarding indication, referral, and reimbursement of the intervention.

The Web-based, diabetes-specific depression treatment seems to have high usability because it can serve as an intervention for both subclinical and clinical depression in diabetes patients. Given its Web-based administration, this diabetes-specific depression treatment has the potential to reach large patient populations and to be cost effective.

Acknowledgments

This study was financed by the Dutch Diabetes Research Foundation (grant number 2005.00.009). The efforts of study participants who gave their time and energy to this work is gratefully acknowledged, as is the excellent work of our coaches and our research assistant Mrs Wilmy Cleijne of the Department of Medical Psychology, VU University Medical Center.

Conflicts of Interest

None declared.

References

1. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004 May;27(5):1047-1053. [Medline: 15111519]
2. Egede LE, Ellis C. Diabetes and depression: global perspectives. Diabetes Res Clin Pract 2010 Mar;87(3):302-312. [doi: 10.1016/j.diabres.2010.01.024] [Medline: 20181405]
3. Schram MT, Baan CA, Pouwer F. Depression and quality of life in patients with diabetes: a systematic review from the European depression in diabetes (EDID) research consortium. Curr Diabetes Rev 2009 May;5(2):112-119. [Medline: 19442096]
4. van der Feltz-Cornelis CM, Nuyen J, Stoop C, Chan J, Jacobson AM, Katon W, et al. Effect of interventions for major depressive disorder and significant depressive symptoms in patients with diabetes mellitus: a systematic review and meta-analysis. Gen Hosp Psychiatry 2010;32(4):380-395. [doi: 10.1016/j.genhosppsych.2010.03.011] [Medline: 20633742]
5. Egede LE. Failure to recognize depression in primary care: issues and challenges. J Gen Intern Med 2007 May;22(5):701-703. [doi: 10.1007/s11606-007-0170-z] [Medline: 17370030]
6. Pouwer F, Beekman AT, Lubach C, Snoek FJ. Nurses' recognition and registration of depression, anxiety and diabetes-specific emotional problems in outpatients with diabetes mellitus. Patient Educ Couns 2006 Feb;60(2):235-240. [doi: 10.1016/j.pec.2005.01.009] [Medline: 16442465]
7. Piette JD, Richardson C, Valenstein M. Addressing the needs of patients with multiple chronic illnesses: the case of diabetes and depression. Am J Manag Care 2004 Feb;10(2 Pt 2):152-162 [FREE Full text] [Medline: 15005508]
8. Snoek FJ. Breaking the barriers to optimal glycaemic control--what physicians need to know from patients' perspectives. Int J Clin Pract Suppl 2002 Jul(129):80-84. [Medline: 12166612]
9. Fisher L, Mullan JT, Arean P, Glasgow RE, Hessler D, Masharani U. Diabetes distress but not clinical depression or depressive symptoms is associated with glycemic control in both cross-sectional and longitudinal analyses. Diabetes Care 2010 Jan;33(1):23-28. [doi: 10.2337/dc09-1238] [Medline: 19837786]
10. van Bastelaar KM, Pouwer F, Geelhoed-Duijvestijn PH, Tack CJ, Bazelmans E, Beekman AT, et al. Diabetes-specific emotional distress mediates the association between depressive symptoms and glycaemic control in Type 1 and Type 2 diabetes. Diabet Med 2010 Jul;27(7):798-803. [doi: 10.1111/j.1464-5491.2010.03025.x] [Medline: 20636961]

11. van Bastelaar K, Cuijpers P, Pouwer F, Riper H, Snoek FJ. Development and reach of a web-based cognitive behavioural therapy programme to reduce symptoms of depression and diabetes-specific distress. Patient Educ Couns 2011 Jul;84(1):49-55. [doi: 10.1016/j.pec.2010.06.013] [Medline: 20619577]
12. Marks IM, Cavanagh K, Gega L. Hands-on Help: Computer-Aided Psychotherapy (Maudsley Monographs No. 49). Hove, UK: Psychology Press; 2007.
13. van Bastelaar KM, Pouwer F, Cuijpers P, Twisk JW, Snoek FJ. Web-based cognitive behavioural therapy (W-CBT) for diabetes patients with co-morbid depression: design of a randomised controlled trial. BMC Psychiatry 2008;8:9 [FREE Full text] [doi: 10.1186/1471-244X-8-9] [Medline: 18284670]
14. van Bastelaar KM, Pouwer F, Cuijpers P, Riper H, Snoek FJ. Web-based depression treatment for type 1 and type 2 diabetic patients: a randomized, controlled trial. Diabetes Care 2011 Feb;34(2):320-325. [doi: 10.2337/dc10-1248] [Medline: 21216855]
15. Andersson G, Cuijpers P. Pros and cons of online cognitive-behavioural therapy. Br J Psychiatry 2008 Oct;193(4):270-271 [FREE Full text] [doi: 10.1192/bjp.bp.108.054080] [Medline: 18827286]
16. National Institute for Health and Clinical Excellence. Depression: Management of Depression in Primary and Secondary Care (amended 2007). Clinical guideline 23. London, UK: NICE; 2004.
17. Grigsby AB, Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. Prevalence of anxiety in adults with diabetes: a systematic review. J Psychosom Res 2002 Dec;53(6):1053-1060. [Medline: 12479986]
18. Peyrot M, Rubin RR. Levels and risks of depression and anxiety symptomatology among diabetic adults. Diabetes Care 1997 Apr;20(4):585-590. [Medline: 9096984]
19. Katon W, Unützer J, Russo J. Major depression: the importance of clinical characteristics and treatment response to prognosis. Depress Anxiety 2010;27(1):19-26. [doi: 10.1002/da.20613] [Medline: 19798766]
20. Fisher L, Skaff MM, Mullan JT, Arean P, Mohr D, Masharani U, et al. Clinical depression versus distress among patients with type 2 diabetes: not just a question of semantics. Diabetes Care 2007 Mar;30(3):542-548. [doi: 10.2337/dc06-1614] [Medline: 17327318]
21. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with Type 2 diabetes. Diabet Med 2008 Sep;25(9):1096-1101. [doi: 10.1111/j.1464-5491.2008.02533.x] [Medline: 19183314]
22. Warmerdam L, van Straten A, Twisk J, Riper H, Cuijpers P. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 2008;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094] [Medline: 19033149]
23. 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] [Medline: 19450912]
24. Beekman AT, Deeg DJ, Van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. Psychol Med 1997 Jan;27(1):231-235. [Medline: 9122304]
25. 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]
26. Snoek FJ, Pouwer F, Welch GW, Polonsky WH. Diabetes-related emotional distress in Dutch and U.S. diabetic patients: cross-cultural validity of the problem areas in diabetes scale. Diabetes Care 2000 Sep;23(9):1305-1309 [FREE Full text] [Medline: 10977023]
27. White IR, Kalaitzaki E, Thompson SG. Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. Stat Med 2011 Nov 30;30(27):3192-3207. [doi: 10.1002/sim.4360] [Medline: 21948462]
28. Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. Br J Psychiatry 2005 Jan;186:11-17 [FREE Full text] [doi: 10.1192/bjp.186.1.11] [Medline: 15630118]
29. Lustman PJ, Griffith LS, Freedland KE, Clouse RE. The course of major depression in diabetes. Gen Hosp Psychiatry 1997 Mar;19(2):138-143. [Medline: 9097068]

Abbreviations

CBT: cognitive behavioral therapy
CES-D: Center for Epidemiologic Studies Depression scale
DbG.nl: Diabetergestemd.nl
DM-distress: diabetes-specific emotional distress
GEE: generalized estimating equation
MDD: major depressive disorder
PAID: Problem Areas in Diabetes scale
RCT: randomized controlled trial

WHO CIDI-auto: computerized version of the World Health Organization Composite International Diagnostic Interview

Edited by G Eysenbach; submitted 29.09.10; peer-reviewed by G Andersson, K Winkley; comments to author 14.03.11; revised version received 13.09.11; accepted 25.09.11; published 05.01.12.

Please cite as:
van Bastelaar KMP, Pouwer F, Cuijpers P, Riper H, Twisk JWR, Snoek FJ
Is a Severe Clinical Profile an Effect Modifier in a Web-Based Depression Treatment for Adults With Type 1 or Type 2 Diabetes? Secondary Analyses From a Randomized Controlled Trial
J Med Internet Res 2012;14(1):e2
URL: http://www.jmir.org/2012/1/e2/
doi:10.2196/jmir.1657
PMID:22262728

Original Paper

Prospective Associations Between Intervention Components and Website Engagement in a Publicly Available Physical Activity Website: The Case of 10,000 Steps Australia

Cally Davies[1], PhD; Kelly Corry[1], BHMSc(Hons); Anetta Van Itallie[1], BHMSc(Hons); Corneel Vandelanotte[1], PhD; Cristina Caperchione[2], PhD; W Kerry Mummery[3], PhD

[1]Centre for Physical Activity Studies, Institute for Health and Social Science Research, CQUniversity Australia, Rockhampton, Queensland, Australia
[2]Faculty of Health and Social Development, University of British Columbia, Kelowna, BC, Canada
[3]Faculty of Physical Education and Recreation, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Cally Davies, PhD
Centre for Physical Activity Studies
Institute for Health and Social Science Research
CQUniversity Australia
Building 18
Bruce Highway
Rockhampton, Queensland, 4702
Australia
Phone: 61 49232263
Fax: 61 49306402
Email: c.davies@cqu.edu.au

Abstract

Background: Effectiveness of and engagement with website-delivered physical activity interventions is moderate at best. Increased exposure to Internet interventions is reported to increase their effectiveness; however, there is a lack of knowledge about which specific intervention elements are able to maintain website engagement.

Objective: To prospectively study the associations of website engagement and exposure to intervention components for a publicly available physical activity website (10,000 Steps Australia).

Methods: Between June and July 2006 a total of 348 members of 10,000 Steps completed a Web-based survey to collect demographic characteristics. Website engagement was subsequently assessed over a 2-year period and included engagement data on website components; individual challenges, team challenges, and virtual walking buddies; and indicators of website engagement (average steps logged, days logging steps, and active users).

Results: On average participants logged steps on 169 (SD 228.25) days. Over a 2-year period this equated to an average of 1.6 logons per week. Binary logistic regression showed that individuals who participated in individual challenges were more likely to achieve an average of 10,000 steps per day (odds ratio [OR] = 2.80, 95% confidence interval [CI] 1.45–5.40), log steps on a higher than average number of days (OR = 6.81, 95% CI 2.87–13.31), and remain an active user (OR = 4.36, 95% CI 2.17–8.71). Additionally, those using virtual walking buddies (OR = 5.83, 95% CI 1.27–26.80) and of older age logged steps on a higher than average number of days. No significant associations were found for team challenges.

Conclusions: Overall engagement with the 10,000 Steps website was high, and the results demonstrate the relative effectiveness of interactive components to enhance website engagement. However, only exposure to the interactive individual challenge feature was positively associated with all website engagement indicators. More research is needed to examine the influence of intervention components on website engagement, as well as the relationship between website engagement and physical activity change.

(J Med Internet Res 2012;14(1):e4) doi:10.2196/jmir.1792

KEYWORDS

Physical activity; engagement

Introduction

Despite the well-documented health benefits associated with engaging in regular physical activity, over half of the people in Western countries are insufficiently active to receive health benefits [1]. Hence, there is a need for effective low-cost physical activity interventions that have a large reach. The Internet is a delivery medium that is able to incorporate interactive, specialized, and individualized health promotion tools to reach large populations at low cost, and is available anywhere and anytime [2,3]. Physical activity promotion programs delivered via the Internet have proven successful in producing short-term behavior change [4-12]. However, it is often reported that these types of interventions have high attrition and limited or declining website engagement [4,10,13-19]; and it is shown that low intervention exposure results in lower overall effectiveness of the intervention [2,16,17,20]. Hence, more research is needed to examine what components enhance website engagement and website use in order to increase their efficiency [14,16,21]. To date, efficacy studies have largely neglected to use objective engagement measures to examine specific intervention components that may increase website engagement [22]. Interactive website features (such as forums, regularly updated content, and online logbooks) have most often been suggested to enhance website engagement; however, there is little evidence to confirm this [22-24].

An existing, freely accessible physical activity promotion program that uses the advantages of Internet delivery to reach a large population is the Australian 10,000 Steps program. The 10,000 Steps program collects advanced website statistics, which allows the opportunity to prospectively examine and identify specific website components associated with website engagement. Studying a publicly available physical activity promotion website offers insight into how people use Internet-delivered programs in real life and can improve the ecological validity of these types of health behavior change programs. This is important, as data from real-life physical activity interventions are scarce, and findings from controlled clinical settings may not effectively translate to real-life settings. Therefore, the aim of the present study was to examine associations between exposure to intervention components and website engagement in a publicly accessible physical activity website (www.10000steps.org.au).

Methods

Participants

The participants were a sample of users of the Step Log feature on the existing 10,000 Steps website (www.10000steps.org.au). In June and July 2006, we randomly selected 663 registered participants from the 10,000 Steps program from a sample of over 24,000 Step Log users. Selected participants were current users, which was defined as having used the Step Log feature for at least 1 day during the month before the survey was conducted. In June and July 2006, we used email to approach potential participants and administer the Web-based demographic questionnaire.

Procedures

The initial email contained an introductory letter outlining the purpose of the research, an invitation to participate, and the URL address, which led to a password protected Web-based questionnaire. By accessing the online questionnaire, participants gave informed consent to be part of the study. To enhance questionnaire completion, participants received three reminder emails to prompt survey completion. The first one was sent 4 days after the initial email, the second at 9 days, and the third 16 days after the original email [25].

Baseline measures of website engagement during May 2006 were extracted from the 10,000 Steps website for all participants who completed the Web-based questionnaire. Additionally, we monitored participants' website engagement over a 2-year period following the completion of the baseline questionnaire. Engagement statistics were downloaded from the 10,000 Steps website 2 years after the baseline data collection (May 2008) for all participants who completed the Web-based questionnaire. A 2-year timeframe was chosen to assess long-term engagement with the website [4-6]. Unique website-user identification numbers allowed for the website statistics, over the 2-year period, to be matched for each individual participating in the survey. Prior to undertaking the study, we obtained ethical approval from the Human Research Ethics Committee at CQUniversity, Rockhampton, Australia.

Intervention

The 10,000 Steps program was initially developed as a multilevel, multistrategy program targeted to the adult population to increase physical activity levels. The program was first delivered in Rockhampton, Australia as a whole-community program based on the socioecological framework [26]. A key aspect of the program is the use of the pedometer to record and monitor physical activity. The use of a pedometer is closely aligned with the prescriptive nature of the program's name, which encourages the accumulation of physical activity in terms of steps per day [26]. Further background on the conceptualization and development of the program has been reported elsewhere [27]. Based on the success of the program [28,29] the funding body, Queensland Health, provided ongoing funding to continue the development, dissemination, and assessment of the program.

This funding supported the development of the 10,000 Steps website, which houses all of the resources and materials for different user groups to implement the 10,000 Steps program in their chosen setting (eg, workplaces). In addition to these resources the website includes an online Step Log (Figure 1), which allows and encourages participants to record and monitor daily physical activity levels in the form of steps (recorded from a pedometer) or time spent in moderate and vigorous physical activity (Figure 2), or both. The 10,000 Steps online Step Log also contains additional interactive components. For example, the individual challenges (I-challenges) allow individuals to choose from a selection of predefined monthly goals (actual steps walked) that correspond to a virtual walking challenge; a new challenge is available each month (Figure 3). Team challenges allow workplaces to offer staff members the opportunity to participate in a virtual walking challenge as part

of a team. Additionally, any registered Step Log member also has the option to add virtual walking buddies, with whom they can share progress through the use of the Step Log. As of July 2011 the 10,000 Steps program had over 160,991 registered participants, who have logged a total of 86,528,244,202 steps.

Figure 1. Screenshot of the online Step Log overview

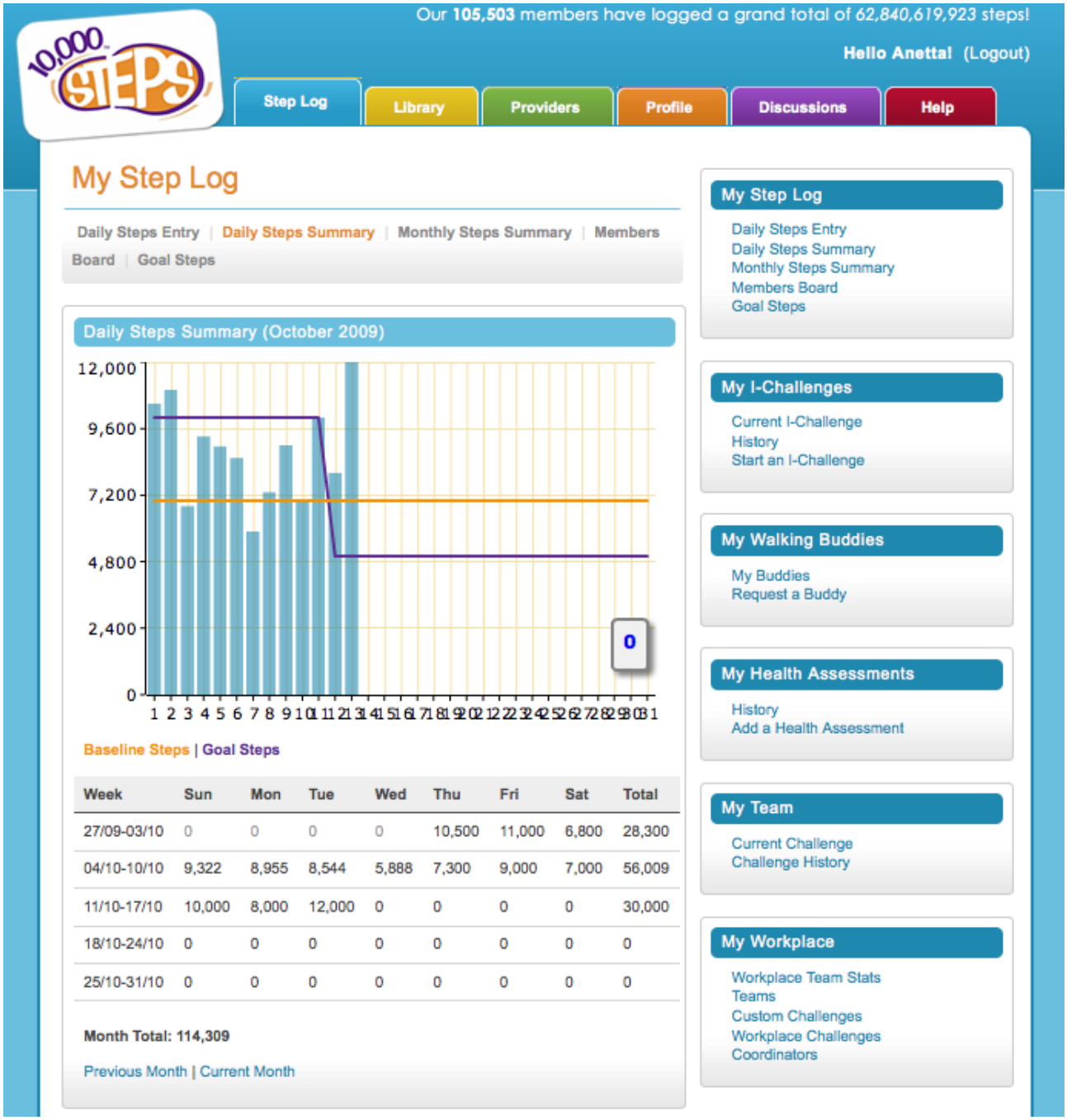


Figure 2. Screenshot of the online Step Log daily steps entry

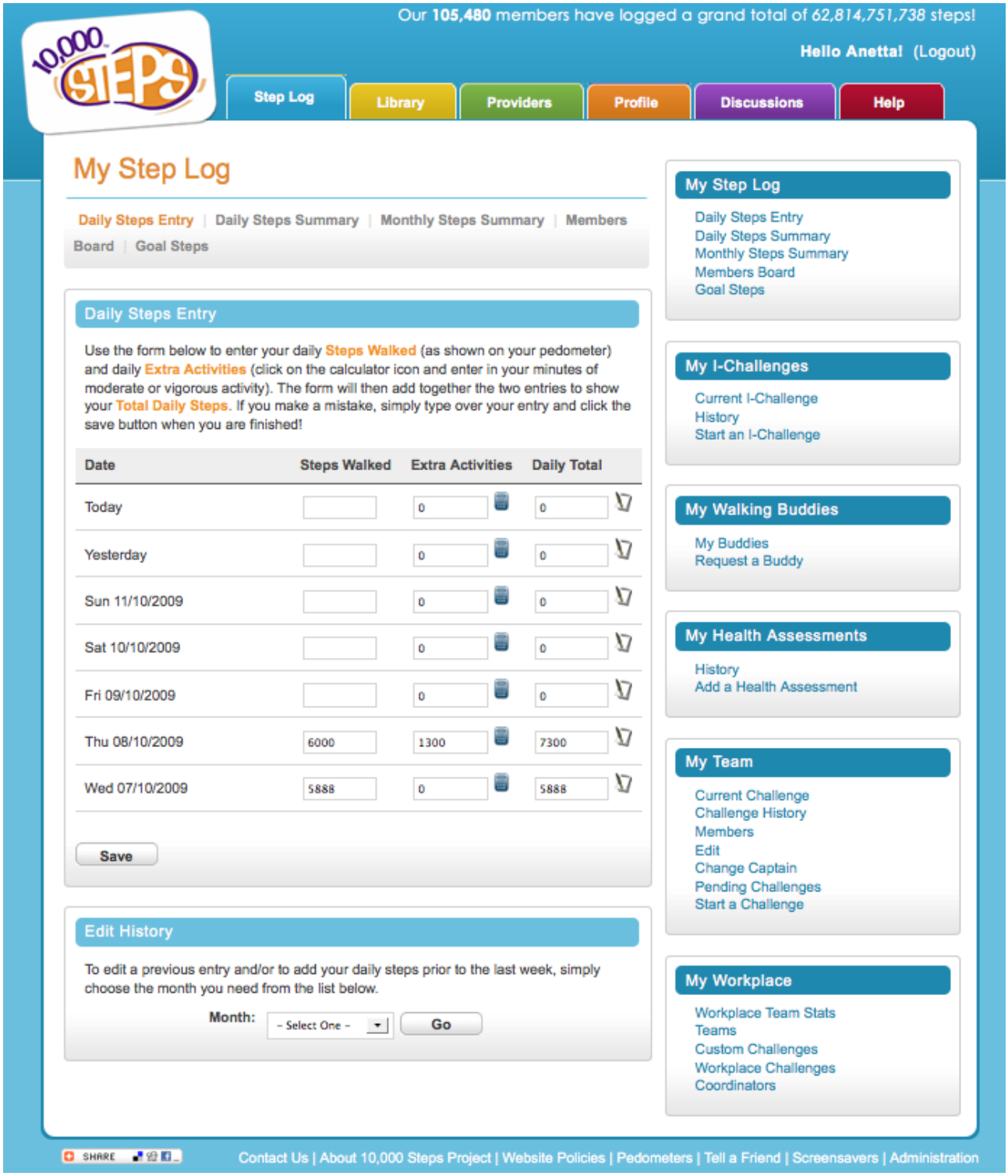
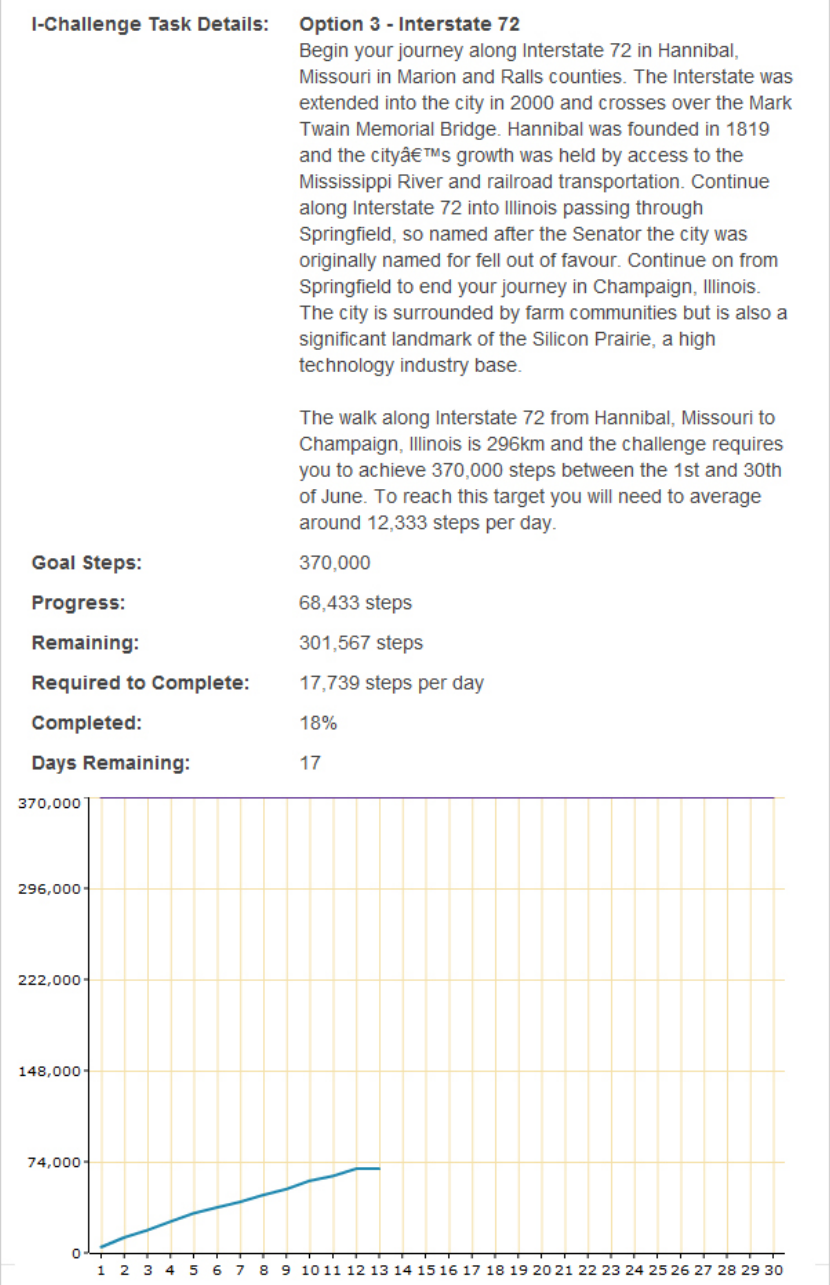


Figure 3. Screenshot of the I-Challenge feature.



Measures

Demographic

The following demographic information was collected from the baseline survey: gender, age category, education level, household income (Australian dollars), body mass index, and presence of a chronic disease. These questions were based on surveys used by the Australian Bureau of Statistics [30].

Website Statistics

Baseline Engagement

As participants in the current study were already using the 10,000 Steps website before the demographic questionnaire was distributed, we extracted 3 baseline engagement measures from the website for May 2006. These measures included the number of months participants had been members of the 10,000 Steps website prior to May 2006; the average number of steps logged on the 10,000 Steps website per Step Log day in May 2006; and the number of days that steps were recorded on the Step Log during May 2006.

Exposure

Exposure to specific website components included use of 3 major interactive features within the 10,000 Steps website (team challenges, I-challenges, and virtual walking buddies). These statistics were used as independent measures to prospectively assess website engagement. Outcomes were dichotomized into a yes-or-no format, indicating the use of the feature on one or more occasion within the 2-year period.

Engagement

Website engagement was expressed through three website engagement variables: average steps, Step Log days, and last Step Log date. These outcome measures were dichotomized to

allow for binary logistic regression as follows: average steps per logged day over the 2-year period was dichotomized into (1) participants who on average recorded 10,000 steps per day or more, and (2) participants who on average recorded less than 10,000 steps per day. The average Step Log days over the 2-year period was dichotomized based on the mean split into (1) participants who equaled or exceeded the mean number of days that steps were logged during the 2-year period across participants, and (2) participants who were below this mean. Finally, the last Step Log date for each participant was used to determine whether participants were still using the website. Participants were dichotomized into (1) users who logged steps on or after January 1, 2008 (active users), and (2) nonusers whose last Step Log date was prior to January 1, 2008.

Website Statistics Rationale

Objective website engagement statistics were measured in 2 categories: exposure to website components and website engagement. Use of the 3 interactive features mentioned above were conceptualized to provide measures for exposure to intervention components. In addition to being 3 of the main components in the 10,000 Steps website, they also facilitate behavior change through theoretical elements such as self-monitoring, goal-setting, feedback, and social support. Therefore, exposure to intervention components was conceptualized as use of the I-challenges, team challenges, or virtual walking buddies over a 2-year period as described below. We hypothesized that exposure to these intervention components would increase website engagement.

We chose 3 separate measures to provide a comprehensive overview of website engagement: average steps per logged day, Step Log days, and the last Step Log date. Each of these measures provides slightly different information about website engagement. First, we conceptualized average steps per logged day as a measure of website engagement; it should not be considered as an outcome measure, as the steps logged on the website did not encompass overall physical activity levels, and we did not measure any other form of physical activity. The Step Log is an intervention feature to benefit participants; it is not a measurement tool. The second measure, Step Log days, is most in line with the norm for measuring website engagement (being number of website logins) but could be considered to be a more in-depth measure, as participants not only log in to the website but also record steps taken. The final measure is last Step Log date. A cut-off date close to the end of the 2-year observational period was chosen (January 2008) to be able to classify participants as current users and therefore currently engaged with the 10,000 Steps program. These 3 forms of website engagement measurements combine to provide an overall view of website engagement and allow for the thorough investigation of associations between exposure to intervention components and website engagement.

Statistical Analyses

We analyzed descriptive statistics of participant characteristics and website engagement. Binary logistic regression analyses were undertaken to prospectively examine the crude and adjusted odds ratios (ORs) for average steps, Step Log days, and last Step Log date. Independent variables assessed included participant demographics and exposure to the specific website components (team challenges, I-challenges, and virtual walking buddies). First, we determined crude ORs for all three dependent variables (average steps logged, Step Log days, and last Step Log date). Next we calculated adjusted ORs for significant crude ORs with the addition of gender, age group, education, income, chronic disease status, number of months of membership prior to May 2006, average steps recorded in May 2006, and number of days steps were logged during May 2006. With the exception of the variables listed, only variables that displayed significant crude ORs were included in the adjusted model; all other variables were excluded, as they did not have a statistically significant effect in the initial model [31]. Statistical analysis was undertaken using SPSS version 17.0 (IBM Corporation, Somers, NY, USA), with the significance level being set to $P < .05$.

Results

Demographics

Of the 663 emails distributed, 15 were undeliverable to addresses used and 300 were not responded to after three prompts, resulting in 348 participants completing the Web-based questionnaires (52.5%). The majority of participants were female (196/306, 64.1%); between 45 and 54 years old (121/306, 39.5%); held a university degree or higher (157/310, 50.7%); and earned under A$75,001per year (157/303, 51.8%). Table 1 presents a descriptive summary of participant characteristics, also broken down by website engagement statistics.

Table 1. Descriptive summary of participant characteristics by website engagement statisticsa

Characteristic			% of average steps >10,000/day		% of Step Log days above average		% of Step Log users in January 2008	
	n	%	n	%	n	%	n	%
Total	348		161	46.3	97	28	66	19
Gender (n = 306)								
Male	110	35.9	63	57	38	35	23	21
Female	196	64.1	83	42	52	27	36	18
Age group (years) (n = 306)								
18–34	62	20	27	44	8	13	10	16
35–44	77	25	33	43	19	25	12	16
45–54	121	39.5	58	48	39	32	23	19
≥55	46	15	29	63	25	54	15	33
Education (n = 310)								
≤ Year 10	29	9	12	41	6	21	4	14
Year 12	37	12	21	57	10	27	7	19
TAFE[b] certificate/diploma or equivalent	87	28	46	53	32	37	20	23
University degree or higher degree	157	50.6	69	44	44	28	30	19
Household income (A$) (n = 303)								
Nil–52,000	93	31	48	52	26	28	16	17
52,001–75,000	64	21	31	48	17	27	15	23
75,001–100,000	42	14	23	55	15	36	7	17
>100,000	62	21	30	48	21	34	13	21
No response	42	14	12	29	11	26	9	21
Body mass index category (kg/m^2) (n = 280)								
<25	110	39.2	55	50	33	30	25	23
25–29.9	106	37.9	55	52	37	35	23	22
≥30	64	23	29	45	16	25	11	17
Chronic condition (n = 310)								
Yes	60	19	27	45	28	47	19	32
No	250	80.6	121	48.4	63	25	42	17

[a] Data collected at baseline (June and July 2006).

[b] Technical and Further Education.

Objective Website Statistics

Baseline Engagement

The mean length of membership before the demographic questionnaire was distributed for the sample was 4.88 (SD 5.02) months. The average number of steps logged per Step Log day during May 2006 was 11,295 (SD 4391.47) and the average days that steps were logged in May 2006 was 22 (SD 11.22).

Exposure

Website engagement data show that the team challenges demonstrated the highest use (137/348, 60.6%), whereas only a minority of participants used the I-challenge or had virtual walking buddies (29.3% and 4.3% respectively; Table 2).

Table 2. Descriptive summary of objective website statisticsa

Characteristic	n	%	% of Average Steps >10,000/day	% of Step Log days above average	% of current users in January 2008
Team challenge (n = 348)					
Yes	211	60.6	45.5	23.2	17.5
No	137	39.4	47.4	35	21.5
Individual challenge (n = 348)					
Yes	102	29.3	60.8	53.9	35.3
No	246	70.7	40.2	17.1	12.2
Virtual walking buddy (n = 348)					
Yes	15	4	60	66.7	33.3
No	333	95.7	45.6	26.1	18.3

[a] Data show website engagement from June 2006 to May 2008.

Engagement

For the indicators of website engagement, average steps per logged day during the 2-year period was 9507 (SD 6665), with 46.3% (161/348) of participants achieving an average of more than 10,000 steps per day for the entire period between 2006 and 2008 (Table 1). Participants reported an average of 169 (SD 228.25) Step Log days, with 28% (97/348) of participants classified as being above the average split. Finally, 19% (66/348) of the participants were still using the 10,000 Steps website in 2008, 20 months after baseline.

Determinants of Website Engagement

The adjusted logistic regression model (Table 3) showed that participants who were male (OR = 1.12, 95% CI 1.13–3.97) or who participated in I-challenges (OR = 2.80, 95% CI 1.45–5.40) were significantly more likely to report average steps above 10,000 per logged day. Participants who were older, participated in I-challenges (OR = 6.18, 95% CI 2.87–13.31), or had virtual walking buddies (OR = 5.83, 95% CI 1.27–26.80) were significantly more likely to log above the average Step Log days. Finally, individuals who participated in an I-challenge (OR = 4.36, 95% CI 2.17–8.71) were significantly more likely to remain active users of the 10,000 Steps website than those who did not.

Table 3. Adjusted odds ratios for website engagement

Characteristic	Adjusted OR[a] (95% CI[b]) average steps >10,000/day	Adjusted OR (95% CI) Step Log days above average (169)	Adjusted OR (95% CI) current users in January 2008
Gender (n = 306)			
Female	1[c]	1	1
Male	1.12 (1.13–3.97)[d]	1.76 (0.84–3.70)	1.23 (0.61–2.45)
Age group (years) (n = 306)			
18–34	1	1	1
35–44	0.96 (0.41–2.25)	2.70 (0.88–8.33)	0.99 (0.37–2.68)
45–54	1.20 (0.54–2.65)	3.02 (1.01–9.04)[d]	1.01 (0.40–2.57)
≥55	2.25 (0.79–6.36)	9.74 (2.75–34.59)[d]	1.88 (0.62–5.71)
Education (n = 310)			
≤ Year 10	1	1	1
Year 12	0.69 (0.24–1.99)	0.50 (0.14–1.85)	0.81 (0.23–2.90)
TAFE[e] certificate/diploma or equivalent	1.59 (0.60–4.18)	0.72 (0.22–2.34)	0.89 (0.30–2.60)
University degree or higher degree	0.85 (0.42–1.70)	1.20 (0.52–2.80)	0.97 (0.45–2.07)
Household Income (A$) (n = 303)			
Nil–52,000	1	1	1
52,001–75,000	0.72 (0.31–1.69)	0.55 (0.19–1.57)	1.61 (0.64–4.08)
75,001–100,000	0.92 (0.36–2.37)	0.95 (0.32–2.88)	0.85 (0.27–2.66)
>100,000	0.66 (0.29–1.54)	0.95 (0.34–2.68)	1.23 (0.47–3.23)
No response	0.42 (0.15–1.14)	0.73 (0.23–2.31)	1.61 (0.56–4.64)
Chronic condition (n = 310)			
No	1	1	1
Yes	0.55 (0.25–1.19)	1.94 (0.80–4.78)	1.44 (0.64–3.24)
Team challenge (n = 348)			
Yes		1	
No		1.07 (0.52–2.19)	
Individual challenge (n = 348)			
No	1	1	1
Yes	2.80 (1.45–5.40)[d]	6.18 (2.87–13.31)[d]	4.36 (2.17–8.71)[d]
Virtual walking buddy (n = 348)			
No		1	
Yes		5.83 (1.27–26.80)[d]	
Baseline engagement			
Length of membership[f]	1.07 (1.01–1.14)[d]	1.21 (1.13–1.29)[d]	1.08 (1.02–1.15)[d]
Average steps[g]	1.00 (1.00–1.00)[d]	1.02 (0.98–1.01)	0.98 (0.99–1.00)
Average Step Log days[g]	0.98 (0.95–1.01)	1.05 (1.01–1.09)[d]	0.99 (0.96–1.02)

[a] Odds ratios (OR) mutually adjusted for all other significant crude OR variables and chronic disease status, gender, age, education, household income, length of membership, average steps in May 2006, and Step Log days in May 2006.

[b] Confidence interval.

[c] Reference category.

[d] Significant OR values; only significant crude OR included and reported in adjusted OR model.

[e] Technical and Further Education.

[f] Length of membership in months prior to May 2006.

[g] Averages observed during May 2006.

Discussion

The primary aim of the present study was to examine associations between exposure to website components (such as interactive features) and website engagement in the Australian 10,000 Steps website (www.10000steps.org.au). This study is unique in that it prospectively examined engagement with a physical activity promotion website over a 2-year period. To our knowledge such an examination has not been published before, despite the demonstrated need to make physical activity promotion websites more effective in terms of behavior change and maintenance [2]. The results show that exposure to interactive intervention components can enhance website engagement in a publically accessible physical activity website in a sample of users who were already actively using the website at baseline. Specifically, exposure to the interactive I-challenge feature was associated with all indicators of website engagement during the 2-year observation period. Additionally, exposure to the virtual walking buddies feature was positively associated with recording above-average Step Log days.

I-challenges are one of the major interactive features incorporated in the 10,000 Steps website; participants choose their own physical activity goal for the month from a set of predefined virtual walking journeys and are able to view their progress through graphs and feedback text. The I-challenge feature was used by only 29.3% of participants, whereas 60.6% used the team challenges feature. Although I-challenges are similar to the team challenges, only exposure to the I-challenge was significantly associated with website engagement. This might be explained by the two prominent differences between these features. Individuals participating in the I-challenges choose to participate entirely of their own accord, whereas individuals participating in a team challenge are recruited by their workplace. This reflects a different motivation to become a 10,000 Steps member and may reflect why participants using the I-challenge feature are more likely to stay engaged with the website [32]. Additionally, new short-term I-challenges are created every month, which allow participants to progressively increase or adjust their goal steps each month. The I-challenge feature also allows individuals to choose from a selection of predefined monthly step goals that cater to different levels of physical activity, whereas the workplace is responsible for setting the overall goal for the team challenges (which in most cases are longer than 1 month). Research highlights the importance of setting personalized goals for ongoing self-management of behavior and to help facilitate behavioral change [33].

The present study reported high levels of program engagement in comparison with previous studies [4,14,21,34], as participants logged an average of 169 (SD 228.25) Step Log days over the 2-year period. This equates to 1.6 logons per week over the 2-year period, which is especially notable, as most previous research has demonstrated lower levels of program engagement over shorter study durations [4,14,21,35]. Marshall et al [35] found that only 26% of participants logged on to their study's website more than once. McKay et al [4] reported an average of 1.1 logins per week throughout their 8 week Internet-delivered physical activity intervention within a sample of individuals with a diagnosis of type 2 diabetes. Steele et al [21] reported an average of 11.8 logins (0.98 logons per week) over the 12-week duration of their physical activity intervention in the general population. Using a longer study design, Lewis et al [36] reported a median of 44 logins (0.86 logons per week) over 12 months for their physical activity intervention. Consistent with the present study, the three studies reporting greater website engagement also incorporated the use of interactive features, which has been suggested to enhance website engagement [5]. This is important, as it has been shown that higher exposure to website content is related to greater behavior change levels [13,37].

Two additional reasons could be associated with the high level of website engagement we observed. First, during the initial 10,000 Steps program (2001–2003), emphasis was placed on marketing 10,000 Steps as a brand [26]. The initial program, a multistrategy community-based intervention, which has been described elsewhere [27], included an overarching marketing campaign. The marketing campaign established high brand recognition and awareness of the 10,000 Steps program that is still being seen today [38]. Since 2004, the 10,000 Steps resources developed in the initial program have been disseminated via the website [26]. By recognizing 10,000 Steps as a brand, participants might be more likely to perceive the website as credible. Website credibility is important for participant engagement, as previous research has shown that users prefer websites that contain credible content [39]. Second, through evaluations of the program, the 10,000 Steps website has been shown to have high levels of usefulness and usability [40], which have previously been shown to be an important factor for increased website engagement [36].

We found that both older age and the presence of virtual walking buddies were related to recording above the average Step Log days. Participants aged 45 years and older were significantly more likely than those aged between 18 and 34 years to engage with the 10,000 Steps program for a longer period of time. The same has been observed in other research [16,18]. Older people may place more importance on their health than younger people do, have more to gain, and are more at risk of developing or already have chronic disease. Additionally, as suggested in previous research, older adults may have more free time at their disposal in contrast to younger adults and as such may be more willing to allocate time to participating in physical activity programs, especially Internet-based programs [41].

The virtual walking buddies feature allows 10,000 Steps members to invite known members to be virtual walking buddies, and they can then view and compare each other's physical activity progress. The positive impact of the virtual walking buddies on website engagement is not surprising, as previous studies have highlighted the importance of social support for behavior change [22,42]. Unfortunately, only 4% of participants made use of the virtual walking buddies in this study. This low number may be a limitation of the 10,000 Steps Step Log, which requires members to know the email address of the person they wish to add as a buddy. More research is needed to explore the effectiveness of virtual walking buddies, as well as how more participants could be encouraged to use it. Virtual social support platforms are becoming increasingly prevalent on the Internet, and exploring the application of virtual social support in relation to health promotion programs is an area of potential future research.

Study Strengths and Limitations

It is important to recognize that there are both limitations and strengths associated with the present study. First, the study incorporates a selective and motivated sample. Included participants are from a sample who self-selected to register for the 10,000 Steps program and who voluntarily elected to respond to the questionnaire. The outcomes of this study might therefore not be entirely comparable with studies that report outcomes under controlled circumstances. Additionally, participants were already engaged with the website prior to the observation period. However, to account for this limitation, length of membership and baseline website engagement were adjusted for in the final model. Furthermore, due to our recruitment method, our study sample might not accurately reflect the demographic characteristics of all users of the 10,000 Steps website. In comparison with previously published statistics for age and gender for all registered members, as of March 2006 the average age was 40.46 (SD 11.74) and 66% were female, compared with an average age of 44.6 (SD 11.74) and 64% female in the current sample [26]. To compare website engagement for the current sample with all users of the 10,000 Steps program, we extracted website engagement statistics for all 10,000 Steps members. Website engagement data for all 10,000 Steps members from the commencement of the website in 2004 to July 2011 demonstrated an average of 12,237 steps per Step Log day and 59.7 Step Log days. Second, the present study did not measure physical activity behavior but instead focused on website engagement. It was not the purpose of the current study to investigate website engagement in relation to physical activity behavior, but rather to examine associations of exposure to interactive website components and website engagement in a physical activity website. Finally, only 19% of participants remained active users of the Step Log in January 2008. This places limitations on the generalizability of the study results to the general population. However, a significant strength of this unique study is the ability to record objective engagement statistics in a publicly accessible (real-life) health promotion website in order to gain a greater understanding of demographic and interactive features that influence website engagement.

Conclusions

This study provides support for the use of a freely available health behavior change program and highlights the importance of including interactive features to enhance program engagement in a sample of individuals already engaged in a Web-based physical activity program. As of July 2011 the 10,000 Steps website had 160,991 registered users, who had logged a total of 86,528,244,202 steps; hence, the potential public health impact of Internet-delivered health promotion programs should not be underestimated. The findings suggest that exposure to interactive intervention components can enhance website engagement. In this particular physical activity promotion program, exposure to the I-challenges and to a lesser extent the virtual walking buddies feature were associated with higher engagement. Engagement in health promotion programs has previously been associated with increased behavior change. In this context higher engagement may lead to increased steps, which are linked to improved health outcomes. Future research should aim to examine the influence of specific intervention components on website engagement, as well as the relationship between website engagement and physical activity behavior change.

Acknowledgments

The findings and conclusions in this paper are those of the authors. The authors would like to acknowledge the continued support of all of the 10,000 Steps members.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2011. Physical Inactivity: A Global Public Health Problem URL: http://www.who.int/dietphysicalactivity/factsheet_inactivity/en/ [accessed 2011-12-15] [WebCite Cache ID 63xTrObKb]
2. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
3. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345. [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]

4. McKay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. Diabetes Care 2001 Aug;24(8):1328-1334. [Medline: 11473065]
5. Napolitano MA, Fotheringham M, Tate D, Sciamanna C, Leslie E, Owen N, et al. Evaluation of an internet-based physical activity intervention: a preliminary investigation. Ann Behav Med 2003;25(2):92-99. [Medline: 12704010]
6. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396 [FREE Full text] [doi: 10.1093/her/cyl096] [Medline: 16971674]
7. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
8. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
9. Morgan PJ, Collins CE, Plotnikoff RC, Cook AT, Berthon B, Mitchell S, et al. Efficacy of a workplace-based weight loss program for overweight male shift workers: the Workplace POWER (Preventing Obesity Without Eating like a Rabbit) randomized controlled trial. Prev Med 2011 May 1;52(5):317-325. [doi: 10.1016/j.ypmed.2011.01.031] [Medline: 21300083]
10. Motl RW, Dlugonski D, Wójcicki TR, McAuley E, Mohr DC. Internet intervention for increasing physical activity in persons with multiple sclerosis. Mult Scler 2011 Jan;17(1):116-128. [doi: 10.1177/1352458510383148] [Medline: 20921239]
11. Glasgow RE, Kurz D, King D, Dickman JM, Faber AJ, Halterman E, et al. Outcomes of minimal and moderate support versions of an internet-based diabetes self-management support program. J Gen Intern Med 2010 Dec;25(12):1315-1322. [doi: 10.1007/s11606-010-1480-0] [Medline: 20714820]
12. Liebreich T, Plotnikoff RC, Courneya KS, Boulé N. Diabetes NetPLAY: A physical activity website and linked email counselling randomized intervention for individuals with type 2 diabetes. Int J Behav Nutr Phys Act 2009;6:18 [FREE Full text] [doi: 10.1186/1479-5868-6-18] [Medline: 19327141]
13. van den Berg MH, Ronday HK, Peeters AJ, Voogt-van der Harst EM, Munneke M, Breedveld FC, et al. Engagement and satisfaction with an Internet-based physical activity intervention in patients with rheumatoid arthritis. Rheumatology (Oxford) 2007 Mar;46(3):545-552 [FREE Full text] [doi: 10.1093/rheumatology/kel341] [Medline: 17043050]
14. Leslie E, Marshall AL, Owen N, Bauman A. Engagement and retention of participants in a physical activity website. Prev Med 2005 Jan;40(1):54-59. [doi: 10.1016/j.ypmed.2004.05.002] [Medline: 15530581]
15. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]
16. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
17. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
18. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer 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]
19. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
20. 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]
21. Steele RM, Mummery WK, Dwyer T. Examination of program exposure across intervention delivery modes: face-to-face versus internet. Int J Behav Nutr Phys Act 2007;4:7 [FREE Full text] [doi: 10.1186/1479-5868-4-7] [Medline: 17352817]
22. 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] [Medline: 21212045]
23. Brouwer W, Oenema A, Crutzen R, De Nooijer J, De Vries N, Brug J. What makes people decide to visit and use an Internet-delivered behavior-change intervention? A qualitative study among adults. Health Education 2009;109(6):460-473. [doi: 10.1108/09654280911001149]
24. Brouwer W, Oenema A, Crutzen R, de Nooijer J, de Vries NK, Brug J. An exploration of factors related to dissemination of and exposure to internet-delivered behavior change interventions aimed at adults: a Delphi study approach. J Med Internet Res 2008;10(2):e10 [FREE Full text] [doi: 10.2196/jmir.956] [Medline: 18417443]
25. Schleyer TK, Forrest JL. Methods for the design and administration of web-based surveys. J Am Med Inform Assoc 2000;7(4):416-425 [FREE Full text] [Medline: 10887169]

26. Mummery WK, Schofield G, Hinchliffe A, Joyner K, Brown W. Dissemination of a community-based physical activity project: the case of 10,000 steps. J Sci Med Sport 2006 Oct;9(5):424-430. [doi: 10.1016/j.jsams.2006.06.015] [Medline: 16890489]
27. Brown WJ, Eakin E, Mummery K, Trost SG. 10,000 Steps Rockhampton: establishing a multi-strategy physical activity promotion project in a community. Health Promot J Austr 2003;14(2):95-100.
28. Schofield G, Steele R, Mummery K, Brown W. Engaging a local council to promote physical activity: the case of dog walking in the 10,000 Steps Rockhampton project. Health Promot J Austr 2004;15(1):78-81.
29. Eakin EG, Brown WJ, Marshall AL, Mummery K, Larsen E. Physical activity promotion in primary care: bridging the gap between research and practice. Am J Prev Med 2004 Nov;27(4):297-303. [doi: 10.1016/j.amepre.2004.07.012] [Medline: 15488359]
30. Australian Bureau of Statistics. Commonwealth of Australia. 2006. 2004-05 National Health Survey: Users' Guide URL: http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/A58E031838C37F81CA257141000F1AAF/$File/4363055001_2004-05.pdf [accessed 2011-12-15] [WebCite Cache ID 63xU68wct]
31. Hoffmann JP. Generalized Linear Models: An Applied Approach. Boston, MA: Pearson Education; 2004.
32. Frederick CM, Morrison C, Manning T. Motivation to participate, exercise affect, and outcome behaviors toward physical activity. Percept Mot Skills 1996 Apr;82(2):691-701. [Medline: 8724948]
33. Funnell MM, Brown TL, Childs BP, Haas LB, Hosey GM, Jensen B, et al. National standards for diabetes self-management education. Diabetes Care 2007 Jun;30(6):1630-1637. [doi: 10.2337/dc07-9923] [Medline: 17526822]
34. Wangberg SC. An Internet-based diabetes self-care intervention tailored to self-efficacy. Health Educ Res 2008 Feb;23(1):170-179 [FREE Full text] [doi: 10.1093/her/cym014] [Medline: 17412717]
35. Marshall AL, Leslie ER, Bauman AE, Marcus BH, Owen N. Print versus website physical activity programs: a randomized trial. Am J Prev Med 2003 Aug;25(2):88-94. [Medline: 12880874]
36. Lewis B, Williams D, Dunsiger S, Sciamanna C, Whiteley J, Napolitano M, et al. User attitudes towards physical activity websites in a randomized controlled trial. Prev Med 2008 Nov;47(5):508-513. [doi: 10.1016/j.ypmed.2008.07.020] [Medline: 18773915]
37. Draper M, Jennings J, Barón A, Erdur O, Shankar L. Research Reports of The Research Consortium of Counseling & Psychological Services in Higher Education. Austin, TX: The Counseling & Mental Health Center, The University of Texas at Austin; 2000. Dose-Effect Relationships in Brief Therapy Based on a Nationwide College Counseling Center Sample URL: http://cmhc.utexas.edu/pdf/DoseResp.pdf [accessed 2011-12-20] [WebCite Cache ID 6448JJhHF]
38. Beehan R, Corry K, Mummery WK. Institute For Health and Social Science Research, Central Queensland University. 2010. 10,000 Steps Working Paper Series, Paper 9: National Awareness of the 10,000 Steps Program, 2009 URL: http://www.10000steps.org.au/pdfs/10K_National_Awareness_Paper_2009.pdf [accessed 2011-12-20] [WebCite Cache ID 6448TENaI]
39. Hinchliffe A, Mummery WK. Applying usability testing techniques to improve a health promotion website. Health Promot J Austr 2008 Apr;19(1):29-35. [Medline: 18481929]
40. Joyner K. Central Queensland University. 2006. 10,000 Steps Working Paper Series, Paper 3: Survey of Current 10,000 Steps Step Log Users URL: http://www.10000steps.org.au/pdfs/workingpapers_paper3_web_000.pdf [accessed 2011-12-20] [WebCite Cache ID 6448U8kbG]
41. van der Mark M, Jonasson J, Svensson M, Linné Y, Rossner S, Lagerros YT. Older members perform better in an internet-based behavioral weight loss program compared to younger members. Obes Facts 2009;2(2):74-79. [doi: 10.1159/000209383] [Medline: 20054209]
42. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]

Abbreviations

CI: confidence interval
I-challenge: individual challenge
OR: odds ratio

Edited by G Eysenbach; submitted 09.03.11; peer-reviewed by L Damschroder, E Anderson-Bill; comments to author 23.04.11; revised version received 14.07.11; accepted 27.09.11; published 11.01.12.

Please cite as:
Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK
Prospective Associations Between Intervention Components and Website Engagement in a Publicly Available Physical Activity Website: The Case of 10,000 Steps Australia
J Med Internet Res 2012;14(1):e4
URL: http://www.jmir.org/2012/1/e4/
doi:10.2196/jmir.1792
PMID:22260810

Original Paper

Primary Care Providers' Perspectives on Online Weight-Loss Programs: A Big Wish List

Kevin O Hwang[1], MPH, MD; Heather L Stuckey[2], EdD; Monica C Chen[1], BA; Jennifer L Kraschnewski[2], MPH, MD; Samuel N Forjuoh[3], MPH, MD, DrPH; Jennifer M Poger[2], MEd; Kathleen M McTigue[4], MS, MPH, MD; Christopher N Sciamanna[2], MPH, MD

[1]Department of Internal Medicine, The University of Texas Medical School at Houston, Houston, TX, United States

[2]Department of Medicine, Penn State College of Medicine, Hershey, PA, United States

[3]Department of Family & Community Medicine, Texas A&M Health Science Center College of Medicine, Scott & White Healthcare, Temple, TX, United States

[4]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Kevin O Hwang, MPH, MD
Department of Internal Medicine
The University of Texas Medical School at Houston
6410 Fannin St
UPB 1100.41
Houston, TX, 77030
United States
Phone: 1 713 500 6441
Fax: 1 713 500 0766
Email: kevin.o.hwang@uth.tmc.edu

Abstract

Background: Integrating online weight-loss programs into the primary care setting could yield substantial public health benefit. Little is known about primary care providers' perspectives on online weight-loss programs.

Objective: To assess primary care providers' perspectives on online weight-loss programs.

Methods: We conducted focus group discussions with providers in family medicine, internal medicine, and combined internal medicine/pediatrics in Texas and Pennsylvania, USA. Open-ended questions addressed their experience with and attitudes toward online weight-loss programs; useful characteristics of existing online weight-loss programs; barriers to referring patients to online weight-loss programs; and preferred characteristics of an ideal online weight-loss program. Transcripts were analyzed with the grounded theory approach to identify major themes.

Results: A total of 44 primary care providers participated in 9 focus groups. The mean age was 45 (SD 9) years. Providers had limited experience with structured online weight-loss programs and were uncertain about their safety and efficacy. They thought motivated, younger patients would be more likely than others to respond to an online weight-loss program. According to primary care providers, an ideal online weight-loss program would provide—at no cost to the patient—a structured curriculum addressing motivation, psychological issues, and problem solving; tools for tracking diet, exercise, and weight loss; and peer support monitored by experts. Primary care providers were interested in receiving reports about patients from the online weight-loss programs, but were concerned about the time required to review and act on the reports.

Conclusions: Primary care providers have high expectations for how online weight-loss programs should deliver services to patients and fit into the clinical workflow. Efforts to integrate online weight-loss programs into the primary care setting should address efficacy and safety of online weight-loss programs in clinic-based populations; acceptable methods of sending reports to primary care providers about their patients' progress; and elimination or reduction of costs to patients.

(J Med Internet Res 2012;14(1):e16) doi:10.2196/jmir.1955

KEYWORDS

Obesity; weight management; primary care

Introduction

One-third of US adults are obese and another third are overweight [1]. The US Preventive Services Task Force recommends that clinicians provide or refer obese adults to high-intensity weight-loss counseling, defined as more than one session per month for at least the first 3 months [2]. Given that many primary care providers lack the time, skills, and supportive infrastructure to provide this level of counseling [3-5], there is an urgent need to identify effective weight-loss resources to which primary care providers can refer their patients.

Online weight-loss programs, with their interactive capabilities and wide reach, have been recognized as potential alternatives to traditional weight-loss programs [6-12]. Establishing partnerships between primary care providers and effective online weight-loss programs could create a substantial public health benefit in which partners play complementary roles to offer the patient a convenient and comprehensive weight-loss service. For example, the primary care provider could identify patients who need to and desire to lose weight, conduct a medical evaluation, and refer eligible patients to the online weight-loss program. Depending on available resources, the online weight-loss program could provide structured counseling, nutrition and exercise monitoring with feedback, and social support [13-16].

Prior studies have assessed patient perspectives on a primary care provider–online weight-loss program partnership [17] as well as clinicians' perspectives on referring patients to weight-loss and diabetes self-education resources [18,19]. Integrating online weight-loss programs into routine primary care will require a thorough understanding of primary care providers' perspectives [5], but such knowledge is lacking. Therefore, a grounded theory approach was used to examine possible theoretical explanations for primary care providers' experiences, attitudes, and preferences with respect to partnerships between primary care providers and online weight-loss programs in routine clinical care.

Methods

Recruitment

We conducted 9 focus group discussions with primary care providers from southeast Texas, central Texas, and central Pennsylvania (Table 1).

Providers were eligible if they practiced general internal medicine or family medicine in the outpatient setting. Participants included physicians, nurse practitioners, and physician assistants. A coordinator at each institution invited potentially eligible primary care providers by email and/or phone and confirmed eligibility.

Table 1. Characteristics of primary care practices for 9 focus group discussions

	Site A				Site B	Site C
Number of focus groups	3				2	4
Practice setting	Urban, academic				Urban, community-based, affiliated with medical school	Suburban, academic
Location	Southeast Texas				Central Texas	Central Pennsylvania
Prevalence of obesity	36.7%[a]				26.1%–31.4%[b]	31.2%[c]
	General internal medicine clinic 1	General internal medicine clinic 2	Family medicine clinic 1	Family medicine clinic 2		
Payer						
Managed care	54%	66%	65%	1%	62%	50%
Medicaid	9%	4%	5%	6%	6%	10%
Medicare	36%	27%	20%	9%	30%	37%
Self-pay, uninsured, or other	1%	2%	11%	84%	2%	3%

[a] Prevalence of obesity among adults (age ≥18 years) seen at general internal medicine and family medicine clinics in 2009–2010. Source: electronic medical records.

[b] Prevalence of obesity among adults (age ≥20 years) in counties served by site B, 2008. Source: Centers for Disease Control and Prevention, National Diabetes Surveillance System.

[c] Prevalence of obesity among adults (age ≥20 years) in county served by site C, 2008. Source: Centers for Disease Control and Prevention, National Diabetes Surveillance System.

Data Collection

The focus groups were facilitated by a general internist with public health training and experience with focus group discussions (KH), a medical student who observed and assisted in leading groups before leading a group (MC), a doctorate-level educator with experience with focus group discussions (HS), and a master's-level educator who observed and assisted in leading groups before leading a group (JP). At the beginning of each session, participants completed an informed consent

form as well as demographic and practice characteristic questionnaires. The facilitators asked open-ended questions to begin the session, using a semistructured standard interview script based on the research objectives. Questions addressed obesity in adults rather than children or adolescents.

The discussions began with introductory questions about how the primary care providers attempted to help their patients lose weight. The current analysis focused on questions related to primary care providers' experience with and attitudes toward online weight-loss programs; useful characteristics of existing online weight-loss programs; barriers to referring patients to online weight-loss programs; and preferred characteristics of an ideal online weight-loss program (Textbox 1). The discussions were recorded and transcribed verbatim. Each focus group lasted approximately 45–60 minutes, and participants received a US $100 gift certificate.

Textbox 1. Focus Group Questions Related to Online Weight-Loss Programs

- If you refer patients to online weight-loss programs, which programs do you use?
- If you had an online and/or in-person resource that you could offer in your clinic to help your patients lose weight that required a minimal time commitment from staff, how interested would your clinic be?
- What characteristics of online weight-loss programs have you found useful for patients?
- What are the top reasons why you don't refer patients to an online weight-loss program?
- What could be offered in an online program that would make you want to refer patients to it?

Data Analysis

The main sources of data were the focus group transcripts, but field notes also included reflections about the focus groups, the settings and culture of the sites, and nonverbal cues during the discussions. Because we found no prior scientific literature on primary care providers' perspectives on online weight-loss programs, there was no well-defined theoretical framework to inform data analysis. Therefore, we used grounded theory to guide the analysis of data [20]. In the grounded theory approach, theory is developed from the data. Features of grounded theory are (1) use of a theoretical sample, (2) constant comparison of data against theoretical categories, and (3) focus on the development of theory via thematic saturation of categories [21]. Three investigators (MC, JK, KH) reviewed the transcripts and field notes, using manual open coding to identify categories, and 3 investigators (JK, KH, HS) convened to discuss common themes within the categories, and compared emergent data against the categories. Disagreements were resolved by consensus. After the ninth focus group, we had the opportunity to conduct an additional group, but we determined that thematic saturation was reached. The study was approved by the institutional review boards of all three sites.

Results

The 9 focus groups included 44 primary care providers with mean age of 45 (SD 9) years (Table 2).

Three major themes from the focus groups were related to barriers to referring patients to online weight-loss programs, while an additional theme identified the characteristics and features of an ideal online weight-loss program (Textbox 2).

Table 2. Characteristics of focus group participants (N = 44)

Characteristic	n	%
Gender, male	25	57
Specialty		
Family medicine	29	66
Internal medicine	12	27
Internal medicine and pediatrics	3	7
Level		
Physician	40	91
Nurse practitioner or physician assistant	4	9
Ethnicity, Hispanic	3	7
Race		
White	32	73
Black	1	2
Asian	10	23
Missing data	1	2

Textbox 2. Major themes from focus group discussions with primary care providers about their perspectives on online weight-loss programs

1. Unfamiliar with online weight-loss programs
2. Uncertain about safety and efficacy of online weight-loss programs
3. Online weight-loss program appropriate only for motivated, technically savvy patients
4. Characteristics and features on an ideal online weight-loss program
   a. Free
   b. Structured curriculum
   c. Goal-setting assistance
   d. Self-monitoring tools
   e. Psychologically oriented content
   f. Peer support
   g. Reports for primary care providers

Primary Care Providers Unfamiliar With Online Weight-Loss Programs

Primary care providers generally reported that they referred their patients to structured weight-loss programs (such as Weight Watchers) or specialists, or they provided counseling within the clinical setting themselves. Many primary care providers had not referred their patients to online weight-loss programs because they were not familiar with them. One said that primary care providers were not educated about online weight-loss programs and that he didn't know of any "online resources to help my patients out with obesity or help them lose weight," and another stated "I'm not aware of them." One participant said:

It's not something that I've routinely done to make specific referrals to online sites and I think as others have said, maybe it's something where I'm just not very knowledgeable of what is available for both professionals and patients.

Some primary care providers had heard of educational websites with obesity-related information, such as MyPyramid.gov, Diabetes.org, and WebMD.com, but were not familiar with "a weight-loss program per se." Others were aware of formal online weight-loss programs such as WeightWatchers.com, SparkPeople.com, CalorieKing.com, FitDay.com, and MyFitnessPal.com, but few had referred patients to them.

Uncertainty About Safety and Efficacy of Online Weight-Loss Programs

Even without being aware of existing online weight-loss programs, the primary care providers expressed uncertainty about the safety and efficacy of online weight-loss programs in general. The main safety concern was that online weight-loss programs would sell unsafe or untested weight-loss medications, either directly or via third-party advertisements. One said, "You'll have a lot of people selling you products that contain unknown chemicals in them and you may make the situation worse."

> *I've had patients come with websites for me to look at where they're being sold something, a stimulant, cathartics, what have you. I think that whatever I'm going to recommend to a patient I have to have gone to myself and look at it. If there's a website where somebody is selling something, that's just not one I would recommend.*
>
> *But that's the biggest concern I have too, that even on a decent site, they are going to be funded. Whoever is funding them is going to set up their little advertisements too. It might be sending the wrong message at times.*

The primary care providers did not specify a degree of weight loss (eg, produce 5% or 10% weight loss) attributed to participation in an online weight-loss program that would increase the likelihood of referring their patients. Acknowledging the paucity of evidence in support of common primary care interventions, a few primary care providers were hesitant to refer patients to online weight-loss programs without evidence of efficacy. One said "I haven't looked at these carefully enough to know...this is one I like and there's good evidence that it works, and I endorse it and suggest you use it."

> *I guess with a lot of the therapeutics that we apply or recommend as physicians, sometimes the evidence base is not so strong and I think we all realize that. So, certainly I guess that if we thought that it wasn't helpful, I guess we might be less inclined to recommend.*

Online Weight-Loss Program Appropriate Only for Motivated, Technically Savvy Patients

Primary care providers thought the online weight-loss programs would be most appropriate for patients who already had skills and self-motivation to lose weight. One provider thought that referring patients to an online weight-loss program would have limited impact because "so many of the people aren't ready to lose weight" and "they just can't get themselves motivated to do so." Another participant thought that online weight-loss programs could be effective when used "in the right place by the right people...in the right frame of mind."

> *Some people need handholding—they really want personal interaction. And then there seems to be a group that are self-starters, that are disciplined—that*

they keep track on their iPhone or program what they eat...They're able to, on their own, make adjustments. The online thing seems to fall kind of in the middle.

If it's online, it could be hit or miss and you would have to have a highly motivated person to keep coming back.

They also thought that older, poorer, or less-educated patients would not or could not access the online weight-loss programs. One provider said "I do have some patients that don't have consistent access or a computer," while another said:

A lot of my patients are Medicaid patients and I don't think they have computer access to begin with. And if they do have computer access, they're using it for recreational purposes.

In summary, most primary care providers were not familiar with online weight-loss programs and they expressed concerns about safety and efficacy. They typically believed online weight-loss programs were most suitable for highly motivated patients who were comfortable with using computers and the Internet.

Characteristics and Features of an Ideal Online Weight-Loss Program

None of the 44 primary care providers claimed to have found an online weight-loss program with all the critical elements that they thought would help patients lose weight. We asked participants to envision the characteristics of a hypothetical, ideal online weight-loss program. The major findings are presented below.

Free

Primary care providers emphasized the importance of patients accessing an online weight-loss program for free. They said they would be more likely to refer patients and that patients would be more likely to join an online weight-loss program that was free.

Structured Curriculum

Primary care providers favored a structured behavioral program with a scheduled curriculum instead of a collection of self-directed resources. Without structure, an online resource would be just like a book: "I don't think that's terribly effective."

It's informal but it's a structured program, allows them to record their caloric intake and caloric expenditure, and gives them some limits that they need to work within depending on what their weight-loss goals are so it's been a nice tool to recommend to people.

If you had a way to generate reminders to people that are visiting the site to say, "hey, did you meet your weight-loss goal this week?," or some type of system so that they don't always have to self‐initiate...People don't want to have a flooded amount of messages from this online weight-loss resource, but it would be kind of nice to know that they're getting reminded...

Goal-Setting Assistance

According to primary care providers, an online weight-loss program should help patients define personal goals. It was also deemed important that users be able to specify "which barrier they want to tackle and how they might choose to do that."

Something that matches the patient's goals, I think, is what's going to be the key. If it's just a series of things that they can do and they're not buying into any [of] them, I don't think they'll be successful. So, I think the motivational part of it has to be what can you see yourself doing moving forward.

They can set a goal weight, so they have a goal that they're shooting for. Then, it interacts with them and gives them a number of calories that they can consume during the day and also then if they exercise, it adds that into the mix.

Self-monitoring Tools

Another feature valued by primary care providers was self-monitoring tools for diet, exercise, and weight. They recognized the opportunity for online weight-loss programs to facilitate the process of self-monitoring of food intake by automatically calculating the calorie content of foods. They felt that the burden of manually entering calories was too high for patients.

[Patients] want something that will kind of show them what they're doing, something that makes it a little bit easier to count their calories.

I'm a firm believer in you got to do a food diary...The way to do it needs to be easy. There needs to be no calculation. There needs to be no nothing. So, like to drink a soda, there's a drop‐down list...It's gotta be easy. Not even saying the calories in it, just let it calculate the calories and give you some analysis later. People need to do no analysis. They just need to report.

Psychologically Oriented Content

Primary care providers felt that an online weight-loss program should offer more than information, that it should also address other mental processes crucial to a successful weight-loss effort. For example, one participant thought it was important that an online weight-loss program address "the motivational aspect of it, and also the implementation...some decision‐making and cognitive informational piece to it." Another provider suggested a problem-solving component: "So, if they don't meet their goal for the week, why did it go wrong? How are they going to get it back on track?" Even straightforward feedback on weight status could be accompanied by psychologically oriented content, such as the following:

...motivation, encouragement, and clearly showing results and benefits to why it's helping you, like showing like in graphs what weight you've lost, how your [body mass index] is changing, how this minimizes your risk factors for heart disease and blood pressure...

Peer Support

Primary care providers recognized the potential value of peer support among users of an online weight-loss program in providing accountability as well as a venue for discussing sensitive issues in "a semi-anonymous fashion." They thought that online peer support could mimic the support from typical group settings. ("Some people would respond to a group setting and so you can obviously do that online.") Connecting patients to other individuals who shared the same struggles would also differentiate an online weight-loss program from less-interactive weight-loss resources, such as books.

> *I think the most important thing is relationships and talking with people, or being accountable to another human being and relating one on one. To the extent that an online program can either simulate a human interaction, or make use of actual people and their experiences, and facilitate experiences through technology, then I think that is an important part. Otherwise, it's like reading a book.*

But the primary care providers were also concerned that online support venues would be a source of "a bunch of bad advice" or "ideas being promulgated as official stuff that's not really correct." One solution would be to have the peer forums monitored by experts.

> *I like having a refereed group where you've got somebody with some education that's chiming in periodically. It's like a group visit in your office, where you've got someone that's educated in that area guiding the group so that if they get off track that you can bring them back. Otherwise I would think that an online discussion would quickly turn into the latest fad.*

Reports for Primary Care Providers

Some primary care providers would welcome reports from the online weight-loss program about their patients' progress or the ability to "check in and see if the patient was using it." The primary care providers anticipated using the reports as a framework for providing praise, support, and accountability either during or between office visits. The reports would position the primary care providers as an accountability partner in the patient's weight-loss effort, because the patient would know "that stuff's going to be going to the physician for review too." The patient would "know you are watching them, instead of them just going off to a website somewhere." This knowledge about the provider's involvement was seen as a motivational factor for patients.

> *I would like to get information. I think if you get it at some kind of pattern, if you're able to respond back to the patient, it makes them accountable, and they might be a little more motivated...They start to worry about their weight, you know, just a few days before the visit, but if they know you're getting things all the time I think it might just be a little more motivation for them.*
>
> *I think to have the actual information available would be a good thing for those people you can call and congratulate or just have a nursing staff just say, "Hey looks like you made progress this month. Congratulations!"*

However, most primary care providers were concerned about the time and effort required to review the reports sent by online weight-loss programs. They thought it would add to their workload, so they preferred "if the website could give feedback to the physician that would not require a great effort on the physician's part to access it." Others thought that feedback from online weight-loss programs should arrive at a controllable frequency, so it would not be overwhelming. Lack of reimbursement was also mentioned as a factor. One provider was reluctant to review and respond to online weight-loss program reports because "currently the reimbursement structure, sad to say, doesn't allow us to do this kind of work." Another said, "Would I want, you know, fifty people telling me to look at their weight program per week and it's not reimbursed? No."

Most primary care providers would want to receive reports only when a patient was not meeting goals because "to get regular progress [reports] on patients who are doing okay is information overload." If a patient is doing well, another participant said, "I don't need all of the detail...it comes down to trying to figure out how much is really enough to trigger some action by us." One participant stated:

> *...it probably would be beneficial to be able to track what they're doing and then help them tweak things if we start to see they're plateauing out on their weight or if they're gaining weight instead of losing weight...*

The primary care providers suggested ways to streamline the communication with online weight-loss programs. The first was to allow the providers to specify the frequency of such reports: "I actually prefer where I'm in control of how often I want to be updated." The second was that online weight-loss programs provide reports electronically and integrate them into existing electronic health records. Providers also had other suggestions to make the online weight-loss program more accessible, such as offering them on mobile devices, on computers located in the clinic, and in other languages (eg, Spanish).

Discussion

Structured online weight-loss programs promote modest weight loss among volunteers from nonclinical settings [6-9] and patients in the primary care setting [10,11]. By exploring the perspectives of primary care providers, we identified core issues to address in translating online weight-loss programs from research settings into routine primary care. This analysis revealed that many primary care providers are not incorporating online weight-loss programs in their patient care. However, the providers raised critical insights about the need for data on program effectiveness and safety; characteristics of patients most likely (in their view) to use online weight-loss programs; program features that providers are likely to endorse; and the types of feedback reports that would facilitate the integration of online weight-loss programs within primary care medicine.

The study had notable strengths. To our knowledge, this is the first qualitative study of clinicians' perspectives on online weight-loss programs. Another strength was the inclusion of primary care providers from multiple practice settings, specialties, and professional designations (physicians and mid-level providers).

The study also had important limitations. The participants were mostly non-Hispanic white or Asian, from urban or suburban practices. Perspectives of providers from other ethnoracial backgrounds and rural settings might have yielded a more complete portrait of the topic. The study did not address the views of other stakeholders, such as patients, office staff, or designers and administrators of online weight-loss programs. Another limitation is that the questions presented in the focus group discussions were not constructed based on a specific theory, nor were they pilot tested before use. We constructed the questions to address clinically relevant gaps in knowledge.

Our results extend knowledge of clinicians' views on referring patients to weight-loss or related resources. A need for better access to such resources has been demonstrated: while 79% of family medicine physicians in New Jersey thought it would be "very helpful" or "crucial" to have a list of community weight-loss resources, only 19% reported knowing "much" or "very much" about community resources for severely obese patients [18]. Likewise, in a national physician survey about diabetes self-management education (DSME) programs, primary care providers noted concerns such as "Do not have enough DSME referral sources" (45%), "Patients are told to do things I do not want" (44%), and "DSME programs do not have quality I want" (31%) [19]. Primary care providers in the current study expressed similar concerns about online weight-loss programs.

According to our study participants, an ideal online weight-loss program would provide a structured curriculum, goal-setting assistance, self-monitoring tools with customized feedback, peer support monitored by experts, and reports for primary care providers. Except for reports for clinicians, these features are common elements of online weight-loss programs [22,23]. Randomized trials have demonstrated the efficacy of such online programs for weight loss[6-9,12,24] as well as maintenance of weight loss [25]. However, the primary care providers in this study did not express awareness of research results or the actual online weight-loss programs used in the trials. This highlights the importance of bolstering efforts to disseminate research and translate interventions into the primary care setting [10,11].

The feasibility and impact of monitoring peer interactions are comparatively less clear. Online peer support holds promise as a useful resource for weight control [14-16], but primary care providers in our study were concerned that some weight-loss advice from online peers would be inaccurate. However, weight-loss advice on online forums has been found to be generally accurate, with medication-related advice more likely than other advice to be inaccurate [26]. Likewise, primary care providers and online forum users were comparable with respect to knowledge about an over-the-counter weight-loss medication, although knowledge in both groups was suboptimal [27]. The effect of expert forum monitoring on weight-loss outcomes remains to be determined.

Primary care providers also preferred that patients have access to online weight-loss programs at no cost. In a prior study, the introduction of out-of-pocket costs for patients reduced participation in and physician referrals to weight-loss and smoking-cessation programs [28]. If patients don't pay, other sources of funding might include insurance carriers, employers, or advertising revenue (content of ads notwithstanding). Automating the counseling would presumably reduce costs. Automated online counseling and human email counseling were both superior to no counseling for weight loss at 3 months [9]. However, wholly automated obesity counseling has been found to be less effective than automated advice augmented with human behavioral email counseling [7]. Regardless of the strategies used to decrease cost, providing effective online weight-loss services at minimal or no cost to patients will require a collaborative effort among multiple stakeholders.

Primary care providers had mixed attitudes about receiving reports from online weight-loss programs, with the desire to track their patients' progress balanced by concerns about time demands. Traditional ancillary providers (eg, physical therapists) send progress reports to referring clinicians. Reports from online weight-loss programs may be necessary if insurance companies were to cover the costs of accessing the programs. However, our results clearly indicate the importance of streamlining the process to minimize the burden on providers in reviewing the reports.

Our findings provide an in-depth view of primary care providers' perspectives on integrating online weight-loss programs into routine clinical care, revealing important areas for research and development as online weight-loss programs continue to be evaluated in clinical populations. The study suggests that efforts are needed to test the feasibility and impact of expert monitoring of peer support forums, develop methods of sending reports to primary care providers about their patients' progress that are acceptable to providers, and minimize costs to patients while providing structured behavioral support. Addressing concerns voiced by primary care providers will hopefully lead to sustainable partnerships with online weight-loss programs, with the end goal of providing patients with comprehensive weight management services.

Acknowledgments

We thank Lynn Medeiros, Dawn Begaye, and Anne Robertson for logistical assistance with this study.

This work was supported in part by a Clinical Investigator Award from The University of Texas Health Science Center at Houston, Center for Clinical Research and Evidence-Based Medicine; the Medical Student Summer Research Program at The University

of Texas Medical School at Houston; and internal funding from Penn State College of Medicine. The funding sources were not involved in reviewing or approving the manuscript for publication.

We presented part of the data as a poster at the 2011 Society of General Internal Medicine annual meeting in Phoenix, AZ.

Conflicts of Interest

None declared.

References

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA 2010 Jan 20;303(3):235-241. [doi: 10.1001/jama.2009.2014] [Medline: 20071471]
2. U.S. Preventive Services Task Force. Screening for obesity in adults: recommendations and rationale. Ann Intern Med 2003 Dec 2;139(11):930-932 [FREE Full text] [Medline: 14644896]
3. Huang J, Yu H, Marin E, Brock S, Carden D, Davis T. Physicians' weight loss counseling in two public hospital primary care clinics. Acad Med 2004 Feb;79(2):156-161. [Medline: 14744717]
4. Yarnall KS, Pollak KI, Østbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? Am J Public Health 2003 Apr;93(4):635-641. [Medline: 12660210]
5. Center for the Advancement of Health. 2001 Apr. Integration of Health Behavior Counseling in Routine Medical Care URL: http://www.prescriptionforhealth.org/downloads/integration2001.pdf [accessed 2011-09-29] [WebCite Cache ID 624DArR3R]
6. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 7;285(9):1172-1177 [FREE Full text] [Medline: 11231746]
7. 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] [Medline: 12684363]
8. Gold BC, Burke S, Pintauro S, Buzzell P, Harvey-Berino J. Weight loss on the web: a pilot study comparing a structured behavioral intervention to a commercial program. Obesity (Silver Spring) 2007 Jan;15(1):155-164 [FREE Full text] [doi: 10.1038/oby.2007.520] [Medline: 17228043]
9. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an Internet weight loss program. Arch Intern Med 2006;166(15):1620-1625 [FREE Full text] [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
10. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
11. McTigue KM, Conroy MB, Hess R, Bryce CL, Fiorillo AB, Fischer GS, et al. Using the internet to translate an evidence-based lifestyle intervention into practice. Telemed J E Health 2009 Nov;15(9):851-858. [doi: 10.1089/tmj.2009.0036] [Medline: 19919191]
12. Harvey-Berino J, West D, Krukowski R, Prewitt E, VanBiervliet A, Ashikaga T, et al. Internet delivered behavioral obesity treatment. Prev Med 2010 Aug;51(2):123-128. [doi: 10.1016/j.ypmed.2010.04.018] [Medline: 20478333]
13. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
14. Krukowski RA, Harvey-Berino J, Ashikaga T, Thomas CS, Micco N. Internet-based weight control: the relationship between web features and weight loss. Telemed J E Health 2008 Oct;14(8):775-782. [doi: 10.1089/tmj.2007.0132] [Medline: 18954247]
15. Trickey AW, Wang W, Hwang KO, Sciamanna CN. Weight loss in a free online weight loss program. 2011 Presented at: 32nd Annual Meeting & Scientific Sessions of the Society of Behavioral Medicine; Apr 27-30, 2011; Washington, DC, USA.
16. Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, et al. Social support in an Internet weight loss community. Int J Med Inform 2010 Jan;79(1):5-13. [doi: 10.1016/j.ijmedinf.2009.10.003] [Medline: 19945338]
17. McTigue KM, Bhargava T, Bryce CL, Conroy M, Fischer GS, Hess R, et al. Patient perspectives on the integration of an intensive online behavioral weight loss intervention into primary care. Patient Educ Couns 2011 May;83(2):261-264. [doi: 10.1016/j.pec.2010.05.009] [Medline: 21459256]
18. Ferrante JM, Piasecki AK, Ohman-Strickland PA, Crabtree BF. Family physicians' practices and attitudes regarding care of extremely obese patients. Obesity (Silver Spring) 2009 Sep;17(9):1710-1716 [FREE Full text] [doi: 10.1038/oby.2009.62] [Medline: 19282824]
19. Peyrot M, Rubin RR, Funnell MM, Siminerio LM. Access to diabetes self-management education: results of national surveys of patients, educators, and physicians. Diabetes Educ 2009;35(2):246-8, 252. [doi: 10.1177/0145721708329546] [Medline: 19208816]

20. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 2nd edition. Thousand Oaks, CA: Sage Publications; 2003.
21. Hood JC. Orthodoxy vs. power: the defining traits of grounded theory. In: Bryant A, Charmaz K, editors. Bryant A, Charmaz K, editors. The SAGE Handbook of Grounded Theory. Los Angeles, CA: Sage Publications Ltd; 2007:151-164.
22. Bensley RJ, Brusk JJ, Rivas J. Key principles in internet-based weight management systems. Am J Health Behav 2010;34(2):206-213. [Medline: 19814600]
23. Rao G, Burke LE, Spring BJ, Ewing LJ, Turk M, Lichtenstein AH, et al; American Heart Association Obesity Committee of the Council on Nutrition, Physical Activity and Metabolism, Council on Clinical Cardiology, Council on Cardiovascular Nursing, Council on the Kidney in Cardiovascular Disease, Stroke Council. New and emerging weight management strategies for busy ambulatory settings: a scientific statement from the American Heart Association endorsed by the Society of Behavioral Medicine. Circulation 2011 Sep 6;124(10):1182-1203. [doi: 10.1161/CIR.0b013e31822b9543] [Medline: 21824925]
24. Micco N, Gold B, Buzzell P, Leonard H, Pintauro S, Harvey-Berino J. Minimal in-person support as an adjunct to internet obesity treatment. Ann Behav Med 2007 Feb;33(1):49-56. [doi: 10.1207/s15324796abm3301_6] [Medline: 17291170]
25. Harvey-Berino J, Pintauro S, Buzzell P, Gold EC. Effect of internet support on the long-term maintenance of weight loss. Obes Res 2004 Feb;12(2):320-329. [doi: 10.1038/oby.2004.40] [Medline: 14981225]
26. Hwang KO, Farheen K, Johnson CW, Thomas EJ, Barnes AS, Bernstam EV. Quality of weight loss advice on internet forums. Am J Med 2007 Jul;120(7):604-609. [doi: 10.1016/j.amjmed.2007.04.017] [Medline: 17602934]
27. Nelson S, Hwang KO, Bernstam EV. Comparing clinician knowledge and online information regarding Alli (Orlistat). Int J Med Inform 2009 Nov;78(11):772-777. [doi: 10.1016/j.ijmedinf.2009.07.003] [Medline: 19716762]
28. Krist AH, Woolf SH, Johnson RE, Rothemich SF, Cunningham TD, Jones RM, et al. Patient costs as a barrier to intensive health behavior counseling. Am J Prev Med 2010 Mar;38(3):344-348. [doi: 10.1016/j.amepre.2009.11.010] [Medline: 20171538]

Abbreviations

DSME: diabetes self-management education

Edited by G Eysenbach; submitted 29.09.11; peer-reviewed by G Bennett, J Johnston; comments to author 26.10.11; revised version received 10.11.11; accepted 17.11.11; published 19.01.12.

Please cite as:
Hwang KO, Stuckey HL, Chen MC, Kraschnewski JL, Forjuoh SN, Poger JM, McTigue KM, Sciamanna CN
Primary Care Providers' Perspectives on Online Weight-Loss Programs: A Big Wish List
J Med Internet Res 2012;14(1):e16
URL: http://www.jmir.org/2012/1/e16/
doi:10.2196/jmir.1955
PMID:22262726

Original Paper

Usability Evaluation of a Web-Based Support System for People With a Schizophrenia Diagnosis

Lian van der Krieke[1], MSc, MA; Ando C Emerencia[2], MSc; Marco Aiello[2], PhD; Sjoerd Sytema[1], PhD

[1]University Center for Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

[2]Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Groningen, Netherlands

Corresponding Author:
Lian van der Krieke, MSc, MA
University Center for Psychiatry
University Medical Center Groningen
University of Groningen
PO Box 30.001
Groningen, 9700 RB
Netherlands
Phone: 31 50 361 21 08
Fax: 31 50 361 97 22
Email: j.a.j.van.der.krieke@umcg.nl

Abstract

Background: Routine Outcome Monitoring (ROM) is a systematic way of assessing service users' health conditions for the purpose of better aiding their care. ROM consists of various measures used to assess a service user's physical, psychological, and social condition. While ROM is becoming increasingly important in the mental health care sector, one of its weaknesses is that ROM is not always sufficiently service user-oriented. First, clinicians tend to concentrate on those ROM results that provide information about clinical symptoms and functioning, whereas it has been suggested that a service user-oriented approach needs to focus on personal recovery. Second, service users have limited access to ROM results and they are often not equipped to interpret them. These problems need to be addressed, as access to resources and the opportunity to share decision making has been indicated as a prerequisite for service users to become a more equal partner in communication with their clinicians. Furthermore, shared decision making has been shown to improve the therapeutic alliance and to lead to better care.

Objective: Our aim is to build a web-based support system which makes ROM results more accessible to service users and to provide them with more concrete and personalized information about their functioning (ie, symptoms, housing, social contacts) that they can use to discuss treatment options with their clinician. In this study, we will report on the usability of the web-based support system for service users with schizophrenia.

Methods: First, we developed a prototype of a web-based support system in a multidisciplinary project team, including end-users. We then conducted a usability study of the support system consisting of (1) a heuristic evaluation, (2) a qualitative evaluation and (3) a quantitative evaluation.

Results: Fifteen service users with a schizophrenia diagnosis and four information and communication technology (ICT) experts participated in the study. The results show that people with a schizophrenia diagnosis were able to use the support system easily. Furthermore, the content of the advice generated by the support system was considered meaningful and supportive.

Conclusions: This study shows that the support system prototype has valuable potential to improve the ROM practice and it is worthwhile to further develop it into a more mature system. Furthermore, the results add to prior research into web applications for people with psychotic disorders, in that it shows that this group of end users can work with web-based and computer-based systems, despite the cognitive problems they experience.

(J Med Internet Res 2012;14(1):e24) doi:10.2196/jmir.1921

KEYWORDS

Schizophrenia; Web-Based systems; Recommendation systems; usability testing; self-management

Introduction

Routine Outcome Monitoring (ROM) has become increasingly important in the mental health care sector. Although there is no universal definition, ROM can be described as the use of standard instruments to systematically and continuously assess aspects of mental health service users' health for the purpose of better aiding their care [1]. The format of ROM varies between countries, but it usually consists of several quantitative measures used to assess a service user's physical, psychological, and social condition. ROM is carried out for service users with a single diagnosis and short-term problems, as well as people with a severe mental illness. This latter group includes service users diagnosed with schizophrenia.

Schizophrenia is a mental disorder characterized by cognitive dysfunctions and abnormalities in perception of reality. People diagnosed with schizophrenia often experience hallucinations, delusions, and disorganized speech and thinking, accompanied by significant social and occupational problems [2]. Due to the complexity of this disorder and the diversity of care needed for service users diagnosed with schizophrenia, proper and frequent evaluation of treatment is particularly vital. That is why ROM offers much potential for better care of these people [3].

However, the effects of ROM on mental health care have been mixed. On the one hand, research shows that the use of outcome measures, combined with adequate feedback, helps clinicians to recognize and anticipate problems in individual treatment processes and to provide better care as a result [4-6]. On the other hand, ROM is not always used to in a way that empowers service users and improves shared decision making between service user and clinician [7,8]. One problem is that clinicians tend to concentrate on those ROM results that provide information about clinical symptoms and functioning. However, service user-oriented approaches promote a focus on personal recovery, which reflects the importance of finding meaning and giving value to personal experiences [7]. A second problem is that service users have limited access to ROM results and they are often not equipped to interpret them [8,9]. These problems need to be addressed, as research has shown that access to resources and the opportunity to share decision making has been indicated as a prerequisite for service users to become a more equal partner in communication with their clinicians [10,11]. Furthermore, shared decision making has been shown to improve the therapeutic alliance, and to lead to better care and treatment [12,13].

Since 2007, ROM assessments have been a regular element in care for people with psychotic disorders in the northern provinces of the Netherlands. The ROM protocol (called PHAMOUS), which is specifically developed for psychotic disorders, consists of a physical investigation (eg, weight, height, waist measurement, and glucose levels), multiple interviews and questionnaires concerning psychiatric and psychosocial issues, and service user satisfaction [14]. All service users with schizophrenia who receive care from any mental health care organization involved take part in ROM assessment at least once a year. After completion of the assessment, the parameters of the ROM assessment are uploaded into a central database by clinicians and research nurses via a link in the patient's service user's electronic file. Currently, the ROM-results are only reported to clinicians. Clinicians are supposed to discuss the results with their patients so that they can mutually decide whether the course of treatment needs readjustment [15]. However, a large percentage of service users do not receive adequate feedback concerning their ROM-results, as clinicians are not yet accustomed to discussing ROM results with service users [16].

In an attempt to improve ROM practice and to increase potential for service user empowerment, we developed a prototype of a web-based support system that provides service users diagnosed with schizophrenia with personalised advice, based on their ROM results. By means of this support system, the current problems with ROM practice may be partly tackled. The personalized advice provides users with accessible information about their ROM results, which may enable them to participate in shared decision making, and pave the way to better care. Prior research has shown that people with psychotic disorders can work with web-based and computer-based systems, despite the severity of their symptoms [eg, 17-21]. Findings are, however, inconsistent as to the amount of support service users need in working with computers (eg, [18] versus [21]).

In the present study, we extended the existing research by investigating the usability of a web-based support system for ROM. We examined whether our support system can make ROM-results more accessible to service users and provide them with more concrete information that they can use to discuss their personal goals with their clinician. The aim of this paper is to provide a brief overview of the web-based system and to report on its usability from the perspective of service users with schizophrenia.

Methods

Content and Technology of the Web-Based Support System Prototype

The prototype of the web-based support system is called WEGWEIS, which is a Dutch abbreviation that stands for web environment for empowerment and individual advice. The WEGWEIS support system offers users advice about various topics related to psychiatric treatment, rehabilitation, and personal recovery. This advice is based on the service user's ROM assessment results, as conducted in the northern provinces of the Netherlands. The support system is a website, which can be accessed by entering a username and a password (see Multimedia Appendix 1). The system is to be used by service users at home or in a clinical setting (eg, a community hospital).

When building the prototype, we focussed on two important and widely used ROM measures, namely the clinician-rated Health of the Nation Outcome Scales [22], which measures health and social functioning, and the service user-rated Manchester Short Assessment of Quality of Life [23], which measures quality of life. Based on item scores of these measures and using innovative algorithms combined with ontological reasoning, the system identifies specific health care problems for each individual service user and provides relevant and

tailored advice [24]. The algorithms are innovative because they break with conventional case-based reasoning approaches in that they decouple symptoms from outcomes, allowing the outcomes to be dynamic [24]. The content of the advice consists of information derived from evidence-based research (eg, the Dutch Multidisciplinary Guideline for Schizophrenia), clinical expertise, and service user experiences.

When, for example, the ROM results indicate that a service user is experiencing physical problems, the system offers advice indicating that physical problems can be a side effect of medication, referring to the Dutch Multidisciplinary guideline for schizophrenia. Furthermore, the advice suggests that side effects may be resolved by adjustment of the medication. Service users are also referred to their psychiatrist – by name – for more information (see Multimedia Appendix 2). When service users appear to experience problems with personal safety, they are provided information about and linked to the local patient counsellor. They also have the opportunity to read about experiences of other service users (see Multimedia Appendix 3). In another example, service users who are troubled by hearing voices are provided a video showing someone suffering from the same condition and offering information about treatment options (see Multimedia Appendix 4). More information about the advice can be found elsewhere [25]. The algorithm for advice selection, as well as a brief overview of system design and architecture are presented elsewhere [24].

The prototype is created with open source software, using the Ruby on Rails Web-framework (http://rubyonrails.org/). The website uses secure connections for all traffic. Service users can access their ROM-results by logging in with a username and password, which are sent to them by email. Failed log-in attempts are logged by the system. ROM-results can only be accessed via patient accounts.

Development of the Prototype

The prototype of the web-based support system was developed by a multidisciplinary team of computer, social, and medical scientists in close collaboration with a group of service users with a schizophrenia spectrum disorder. The content and functionality of the first prototype was based on a needs assessment (unpublished material) conducted in 2009, consisting of semi-structured interviews with service users, relatives of service users, nurses, psychologists, psychiatrists, and people involved in e-mental health services for people with a psychiatric disability.

We put particular focus on the design of the support system's user interface, as it has been suggested that people with schizophrenia have special needs with regard to web design [26]. This is supported by the theory that the quality of a user interface is partly determined by the extent to which users are able to create a so-called mental model of the website. A mental model can be described as a representation of a person's thought processes regarding the functionality and structure of the website, and the flow of information therein. Therefore, it is important for designers to match as closely as possible the user interface with this mental model [27]. Finding a good match can be particularly challenging. This is especially the case when dealing with people with schizophrenia, who experience cognitive problems such as concentration, memory and information processing difficulties [26]. As a result, their mental models may differ from those of other users.

A few studies have investigated the challenges in web design for people with a schizophrenia diagnosis. Results from these studies suggest that users with schizophrenia experience difficulties with stimulus overflow, large amounts of text or information, interpretation of two-word labels, and remembering previous steps in the navigation process [17,18,26,28]. Furthermore, some of them experience paranoia when using computers and Internet [17].

In conjunction with the general guidelines as described in *User Interfaces for all* (a handbook for user interface design) [29] and taking into account the findings from prior research, we set out some specific rules for the design of the support system's interface. The most important of these specific rules were the following: no use of unexpected pop-ups, transparency of procedures (ie, clear information about what happens when users click a button, what purposes their personal information is used for and who it is available to, etc), use of concrete descriptions (including using the name of a service user's psychiatrist, instead of the general designation 'your psychiatrist'), limited amount of text on one screen with an option to increase/decrease the amount of information, use of video material in addition to text, limited number of bright colours and avoiding jargon or difficult terms.

Participants

Service users were recruited from four mental health care organizations in the Netherlands through snowball sampling. Snowball sampling involves asking a key informant or study participant whether they can suggest a person who fits the study criteria and asking them to introduce this person to the researcher [30]. In our case, study participants were recruited by 5 clinicians and fellow study participants. The study was conducted in March and April 2011. The inclusion criteria were (1) having a diagnosis of schizophrenia or a related psychotic disorder (eg, schizo-affective disorder, schizophreniform disorder, schizotypal disorder), (2) being between 18 and 65 years old and (3) being fluent in Dutch. There were no exclusion criteria.

Sixteen service users were asked to participate and a total of 15 service users, 10 male and 5 female, agreed to participate in the study. The age of the participating service users ranged from 23 to 61 years, with a mean age of 42. The duration of illness for 13 of these service users was known and ranged from 3 to 25 years, with a mean duration of 13 years. All service users received care in an outpatient setting except for one, who was committed in a forensic setting. In order to provide participants with some time to consider their participation, they were informed about the purpose and content of the testing by either a clinician or one of the experimenters (LvdK) at least a week prior to testing. Directly before the usability testing was to start, written informed consent was obtained. After completing the study, participants received a gift voucher of 15 euros.

Four Information and Communication Technology (ICT) experts participated in the study. They fulfilled the role of evaluator in

a heuristic evaluation process, as described below. All ICT experts were experienced in developing ICT applications for mental health care organisations.

Usability Testing

Usability can be defined as the ease with which users can use a particular tool or object to achieve a specific goal. One of the leading experts on usability, Jakob Nielsen, distinguishes five main quality components of usability [31]: (1) *learnability*: how easy is it for users to accomplish basic tasks the first time they encounter the design; (2) *efficiency*: once users have learned the design, how quickly can they perform tasks; (3) *memorability*: when users return to the design after a period of not using it, how easily can they re-establish proficiency; (4) *errors*: how many errors do users make, how severe are these errors, and how easily can they recover from the errors; and (5) *satisfaction*: how pleasant is it to use the design.

Usability can be assessed by usability testing. There are three testing categories: heuristic evaluation, qualitative evaluation, and quantitative evaluation. These categories will be described in the following sections.

Heuristic Evaluation

We started the usability testing by conducting a heuristic evaluation. This is a research method for detecting usability problems with the interface early in the testing process [31]. Heuristic evaluation is conducted by evaluators and takes place prior to the testing by end-users (in our case service users). Problems detected by the evaluators are dealt with immediately so they do not influence the rest of the testing process.

Heuristic evaluation is usually conducted by more than one evaluator because it is difficult for one person to detect all usability problems. We appointed four ICT experts to fulfill the role of the evaluator, as this falls into the range of the optimal number [32]. The process of heuristic evaluation used in this study is based on Nielsen's recommendations [33]. The evaluators were given a brief introduction to the background and rationale of the web application under review, then given instructions on how to conduct the heuristic evaluation. One of the most important instructions was that they were not allowed to communicate with each other during the testing process. Then, the evaluators sat at the computer and went through the user interface according to a scenario written by the experimenters. The scenario included using log-in procedures, username and password retrieval processes, font size modification, completing questions, going through advice units, printing information, searching for advice by means of key words, and providing feedback about the website. The evaluators inspected the interface independently, assessing the various elements based on a list of ten recognized usability principles ("heuristics") [33] translated into a series of questions (see Table 1). Their findings were put in a template developed by the experimenters.

The data in the four completed templates was assembled in one document and its content was analysed, meaning that the data was categorized according to Nielsen's usability topics [33] (see also Table 1). Finally, a list of usability violations was created and sorted according to frequency and priority. A debriefing meeting was organized with evaluators and the experiments, during which the results of the heuristic evaluation were discussed during a brainstorm session in a brainstorm mode. Decisions were made as to which usability issues were considered most urgent and how these issues could best be solved.

Table 1. Assessment criteria for heuristic evaluation

Usability principle	Question
1. Visibility of system status	Are there any incidents where the website is unresponsive or slow?
2. Match between system and the real world	Are there any words/sentences used on the website that do not match the language used by the intended group of users?
3. User control and freedom	Are there any instances where important changes made by users cannot be easily undone?
4. Consistency and standards	Are there any inconsistencies concerning language use or functionality?
5. Error prevention	Are there any instances where users can easily make mistakes? Before executing an action, are users asked for confirmation where needed?
6. Recognition rather than recall	Are there any pages where the content or structure is unclear or insufficiently explained?
7. Flexibility and efficiency of use	Are there any frequently used functionalities on the website that are not accessible fast enough?
8. Aesthetic and minimalist design	Are there any instances in which the website offers too much information, whereby the user can loose track of the situation?
9. Help users recognize, diagnose, and recover from errors	Are there any error alerts which are not clear to users, which do not identify the problem correctly or do not provide a solution?
10. Help and documentation	Is there enough help or documentation available?

Qualitative Evaluation

After completion of the heuristic evaluation, we conducted a qualitative evaluation. In this process, end-users fulfilled the role of the evaluator. The participants were invited to sit at a computer. We then asked them to use the web application following a scenario written by the experimenters (the same scenario as used in the heuristic evaluation). Users were encouraged to work through the scenario step by step, starting with the log-in procedures. We decided not to ask participants

to think aloud, as we suspected that this might affect their way of working substantially. Two-thirds of the end-user participants carried out the testing at our research centre. During the testing, one of the experimenters observed the users' actions via a beamer projection on a screen, while making notes (see Multimedia Appendix 5). One-third of the users conducted the testing at home on their own computer and were joined by an experimenter who observed from a distance. When users finished the testing, they were asked to verbally describe their first impression of the support system. As the main aim of this part of the testing was to find out how users interact with the web system, the research method used in this qualitative evaluation was (non-participant) observation [34]. One experimenter (LvdK) was present during the testing session and made notes (using paper and pencil) which indicated how participants worked their way through the scenario. The sessions were not audiotaped, as observation was the main evaluation method and we assumed that participants might not feel at ease with audiotaping. The verbal information provided by service users was analysed by identifying positive and negative feedback items.

Quantitative Evaluation

After the qualitative evaluation was completed, a quantitative evaluation was conducted. End-user participants were asked to fill out a short questionnaire, consisting of 5 questions measured on a 5-point Likert scale. They were asked about their computer and Internet use. This questionnaire was derived from another European study testing a web application developed for a comparable group of end-users [18]. Furthermore, participants completed a Satisfaction Questionnaire, measuring their satisfaction with various aspects of the web application concerning layout, structure, user-friendliness and content. This questionnaire consisted of 13 statements to be subsequently rated on a 7-point Likert scale, ranging from completely disagree (1) to completely agree (7). The Satisfaction Questionnaire was specifically designed for this study by the research group. Descriptive analysis (mean, standard deviation) of the quantitative data was conducted with SPSS 16.0 statistical software for Windows (SPSS Inc., Chicago, IL, USA).

Results

The results of the usability tests are a combination of the three categories of testing mentioned above, namely heuristic evaluation, qualitative evaluation, and quantitative evaluation.

Heuristic Evaluation

All ICT experts evaluating the website were able to complete the scenario written by the experimenters. No major problems were reported with regards to language, undoing changes, structure or content of the pages, accessibility of functionality and clarity of error messages (ie, usability principles 2, 3, 4, 6, 7 and 9). However, there were some instances in which the website was unresponsive or slow. Furthermore, at times the website seemed to offer too much information at once, and three situations occurred whereby users were not clearly directed to the right page. The most obvious problem reported was that the Disclaimer page was empty and that there was no existing Help section or Frequently Asked Questions section.

During the debriefing meeting, all problems were discussed and decisions were made on how to solve problems most effectively. All problems were solved prior to the qualitative and quantitative testing with service users, except for the missing Frequently Asked Questions section, which was composed after the usability testing with service users.

Qualitative Evaluation

All end-user participants were able to complete the scenario, although three of them needed some hints in order to continue to the next step. For instance, one participant had difficulty finding out how to adjust his personal profile, and the experimenter had to explain how he could access the profile. Although the participants were not asked to think aloud during the evaluation, most of them did so spontaneously. One of the difficulties expressed was that some buttons were hard to find or that their function was not entirely clear. One example is the ‘Feedback’ button. This button was located at the left part of the webpage, situated vertically and separately from the navigation bar. Three participants could not immediately locate it and two did not know what to use it for. Furthermore, several participants suggested that the website could be made more attractive by using more colour, more images and videos, and more links. However, others indicated they were happy with the layout and found the website to be nice and simple.

With reference to the content of the website, participants expressed that they recognized many issues that people suffering from schizophrenia are faced with and believed that the website could be a useful instrument in supporting people in their personal recovery process. In addition, while reading the advice, various service users came up with relevant information that they thought should be added to the advice. A few other participants, however, stated that the information about illness symptoms and medication should be more extensive. In addition, one participant suggested creating a possibility for online communication between clinicians and service users within the system.

Quantitative Evaluation

The participating end-users reported to be well experienced in using computers and the Internet, to have good computer and Internet skills (see Table 2) and to have a positive attitude towards technology (see Table 2). There was one participant who reported to have almost never used the Internet. He appeared not to have access to the Internet, due to the fact that he was a forensic service user admitted into a penitentiary where Internet use was not allowed.

The mean score of satisfaction with the web-based support system prototype was 73.60 (the maximum being 90) with a standard deviation of 6.64. Ratings of the individual statements are presented in Table 3. As this table shows, the most disagreement amongst the participants concerned the question of whether or not the website was boring. This is in line with the results of the qualitative analysis, which showed that some participants found the website nice and quite simple, whereas others suggested that it could be improved by using more colour, images, and so on.

Table 2. Service users' computer and Internet use, skills and attitudes

	Questionnaire response option	No. (n=15)
Computer use	Almost never	0
	Less than once a month	0
	Monthly	0
	Every week	1
	Every day	14
Internet Use	Almost never	1
	Less than once a month	0
	Monthly	0
	Every week	1
	Every day	13
Computer Skills	Very bad	1
	Bad	0
	Not bad, not good	5
	Good	8
	Very good	1
Internet Skills	Very bad	1
	Bad	0
	Not bad, not good	4
	Good	9
	Very good	1
Attitude towards computers	Very negative	0
	Negative	0
	Neutral	0
	Positive	11
	Very positive	4

Table 3. Results of Satisfaction Questionnaire

	Mean (SD)	Percentage (and absolute number) of service users who agreed (score 6) or completely agreed (score 7) with the statement (N = 15)
I can easily find my way on the website	5.73 (0.88)	80 (12)
I am satisfied with the language used on the website	6.13 (0.35)	100 (15)
The website is boring	3.13 (1.55)	7 (1)
I am satisfied with the font used on the website	5.87 (0.83)	93 (14)
The colour of the website was appealing	5.33 (1.35)	67 (10)
The website does not contain distracting elements	5.8 (1.21)	80 (12)
The advice provides me with meaningful information	5.67 (0.72)	80 (12)
The amount of information in the advice is too much	2.87 (1.55)	7 (1)
The advice can help me reflect on what I want	5.73 (1.16)	80 (12)
I can imagine myself discussing the advice with my clinician in the future	5.67 (1.11)	80 (12)
I can imagine the advice being helpful to others	6.27 (0.46)	100 (15)
I think I will use the website in the future	5.53 (0.83)	60 (9)
I would recommend the website to others	5.87 (0.64)	86 (13)

Discussion

In this study, we investigated the usability of the first prototype of a web-based support system for people diagnosed with schizophrenia. The heuristic evaluation with ICT experts revealed some minor problems; the most important ones of which were unresponsive or slow processing of information, too much information displayed at once, an empty Disclaimer page and no existing Help section. The first three problems were solved before testing with service users. During qualitative testing, our group of end-users reported some difficulties with, among other things, the location and function of the 'Feedback' button and with understanding how to adjust one's personal profile. In addition, several suggestions were made to make the interface more attractive. These results indicate that the end-users involved in this study, varying in age, sex and duration of illness, were able to use the support system easily. Furthermore, the content of the advice generated by the support system was judged to be meaningful and supportive. We can therefore conclude that, overall, the support prototype has valuable potential for improving the ROM practice and that it is worthwhile to develop it further into a more mature system.

Comparison With Existing Research

Our preliminary results are in line with previous research, which shows that people with psychotic disorders can work with web-based and computer-based systems [17-21], but there are some differences between our research and that of others that we need to address.

Whilst designing the interface, we followed some specific rules based on existing literature in the field and for this group of end-users as well as applying general rules of interface design. However, we did not comply with all recommendations presented in the literature as feedback from individual service users during the design process, which took place prior to the usability testing (not described in this paper), suggested it might not be necessary. For instance, we decided to use a bright background colour (yellow) for the web pages, and we used arrow heads and drop down menus instead of pop-ups, which was advised against by Rotondi et al, [20]. However, these deviations did not result in any usability violations.

This may be explained by the fact that there appears to be a difference between basic principles for user interface design and concrete applications thereof. Each basic principle can be translated into various concrete applications. If the principle is to avoid an abundance of information, this can be achieved by either limiting the amount of text on one page, or by ordering the information in a surveyable way. Both forms can be effective, depending on, among other things, users' individual preferences. Furthermore, as the functionality of Internet browsers develops very quickly and new innovations emerge, some earlier problems with the user interface may be no longer relevant. For instance, Rotondi et al [20] discourage the use of an absolute font size that cannot be enlarged. Given the flexibility of modern-day browsers, however, this is hardly an issue anymore, as font sizes can be adjusted rather easily.

Another issue to be addressed is the context for which the support system is developed. As mentioned before, our system is intended for independent use by service users at their home or on a hospital ward. This is in line with the study by Bickmore et al [21], who developed a computer-based medication adherence system with relational agents for service users with schizophrenia, to be used at home and without assistance or interpretation from clinicians. Results of their pilot evaluation study (N = 16) showed that independent use of the computer system was acceptable for all but one of the study participants, who were recruited at an outpatient clinic. However, these results seem to contradict with the findings of Kuosmanen et al [18], who reported that service users with psychotic symptoms needed support from nurses in using their web system. This

difference in findings could be explained by symptom severity of service users, as the study by Kuosmanen et al [18] was conducted in a locked-door setting, while the one by Bickmore et al [21] and our study primarily involved service users staying at home.

The results of our study add to previous studies in that usability tests suggest that there need not be insurmountable barriers in independent use of web-based systems for people with psychotic disorders. However, we need to investigate the system in a real world setting in order to draw broader conclusions. In future research, the most important question will be not so much whether or not service users with psychotic symptoms can independently work with web systems, but rather, under what conditions they can successfully work with them. These conditions may depend upon the service users' circumstances, such as receiving care in an inpatient or outpatient setting, severity of specific symptoms (eg, paranoid ideas), and, of course, the level of computer experience. In addition, they might also be related to the web-system, such as the content and the complexity of the system's functionality.

Future Development of Our Web-Based Support System

In future development of our web-based support system, several issues need to be taken into account. First, in order to provide end-users with a support system tailored to their needs and preferences, a flexible interface is needed. Some users like a colourful background and all kinds of multimedia elements, whereas others prefer a more simple interface. This calls for an interface which can be customised. Furthermore, in order to keep the content of the advice oriented toward the service user and to work in a more Health 2.0/Medicine 2.0 fashion [35], we need to facilitate users in adding information to the advice by creating options to post comments or upload material. In addition, possibilities should be explored for interactive communication among the service users themselves and between the service users and clinicians.

With regard to the support system's technology, we aim to develop more sophisticated advice algorithms and enlarge our data set so that the advice offered to service users can be even more personalised. Furthermore, we will explore interoperability and connectivity with personal health records and electronic patient files, and integration with successful platforms currently used in mental health care.

Limitations

Our study should be viewed with consideration of certain limitations that we encountered. First, our sample of service users was small and we used a method of snowball sampling, which is a form of convenience sampling. One disadvantage of convenience sampling is that one runs the risk of compiling a non-representative study sample. In our case, the study sample was quite diverse in age, sex, and duration of illness, which favours the sample's representativeness.

In contrast, what appears to be less favourable for the sample's representativeness is the fact that the service users recruited for this study might have had a particular interest in working with computers and websites, which could have affected our results. This could be the case given that the service users concerned were reported to be quite skilled in using the computer and Internet. However, we need to take into account that the Netherlands is one of the countries with the highest Internet penetration rates. In March 2011, 88.3% of the Dutch population had Internet access, while the world wide average is only 30.2% [36]. This suggests skillful computer and Internet use is not uncommon in the Netherlands. Understandably, there will be differences between the level of computer and Internet skills of the general Dutch population and people with mental disorders. However, we believe that the representativeness of our sample on this point does not necessarily invalidate our conclusions.

Second, the presence of an experimenter during the testing session may have affected the behaviour of service users conducting the testing. Although the experimenter encouraged participants to mention both strong and weak features of the web application, they might have felt reluctant to be critical.

Third, the support system was not tested in the context of a full ROM assessment, but as a somewhat isolated part thereof. Therefore, at the moment, we cannot gain a comprehensive view of the system's functioning in its full setting. This issue needs to be addressed in future research in a clinical evaluation, followed by an examination of its effectiveness in a randomized controlled trial, in order to determine whether or not the present system can genuinely contribute to improving ROM practice.

Acknowledgments

This study was supported by the Netherlands Organisation for Health Research and Development (ZonMw); Fonds Psychische Gezondheid; ICT regie; and the Dutch Ministry of Health, Welfare and Sport under the name WEGWEIS (grant number 300020011). More information can be found on the project's website: http://development.wegweis.nl. The authors thank Prof. dr. D. Wiersma and Prof. dr. N. Petkov for their valuable contribution to the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Homepage Wegweis website.

[PDF File (Adobe PDF File), 69KB - jmir_v14i1e24_app1.pdf]

Multimedia Appendix 2

Advice can refer to the Dutch Multidisciplinary Guideline and encourage service users to contact their psychiatrist.

[PDF File (Adobe PDF File), 67KB - jmir_v14i1e24_app2.pdf]

Multimedia Appendix 3

Users have the opportunity to learn about the experiences of other service users.

[PDF File (Adobe PDF File), 67KB - jmir_v14i1e24_app3.pdf]

Multimedia Appendix 4

Service users can watch a video of someone who experiences hearing voices.

[PDF File (Adobe PDF File), 303KB - jmir_v14i1e24_app4.pdf]

Multimedia Appendix 5

Test situation at our research center. During the testing, one of the experimenters observed the users' actions via a projection on a screen while making notes.

[JPG File, 575KB - jmir_v14i1e24_app5.JPG]

References

1. Trauer T. Introduction. In: Trauer T, editor. Outcome measurement in mental health: theory and practice. Cambridge: Cambridge University Press; 2010:1-14.
2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV. Washington, DC: American Psychiatric Association; 2000.
3. Opler LA, Ramirez PM, Mougios VM. Outcome measurement in serious mental illness. In: Ishak WW, Burt T, Sederer LI, editors. Outcome measurement in psychiatry: a critical review. Washington, DC: American Psychiatric Pub.; 2002:139-154.
4. Lambert MJ, Hansen NB, Finch AE. Patient-focused research: using patient outcome data to enhance treatment effects. J Consult Clin Psychol 2001 Apr;69(2):159-172. [Medline: 11393594]
5. 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] [Medline: 15609358]
6. Whipple JL, Lambert MJ. Outcome measures for practice. Annu Rev Clin Psychol 2011 Apr;7:87-111. [doi: 10.1146/annurev-clinpsy-040510-143938] [Medline: 21166536]
7. Lakeman R. Standardized routine outcome measurement: pot holes in the road to recovery. Int J Ment Health Nurs 2004 Dec;13(4):210-215. [doi: 10.1111/j.1445-8330.2004.00336.x] [Medline: 15660588]
8. Guthrie D, McIntosh M, Callaly T, Trauer T, Coombs T. Consumer attitudes towards the use of routine outcome measures in a public mental health service: a consumer-driven study. Int J Ment Health Nurs 2008 Apr;17(2):92-97. [doi: 10.1111/j.1447-0349.2008.00516.x] [Medline: 18307597]
9. Happell B. Meaningful information or a bureaucratic exercise? Exploring the value of routine outcome measurement in mental health. Issues Ment Health Nurs 2008 Oct;29(10):1098-1114. [doi: 10.1080/01612840802319852] [Medline: 18853349]
10. GGZ Nederland. Naar herstel en gelijkwaardig burgerschap. Visie op de (langdurende) zorg aan mensen met ernstige psychische aandoeningen. URL: http://www.ggznederland.nl/scrivo/asset.php?id=305955 [accessed 2011-12-09] [WebCite Cache ID 63ndB6FmG]
11. Deegan PE. Recovery and empowerment for people with psychiatric disabilities. Soc Work Health Care 1997;25(3):11-24. [doi: 10.1300/J010v25n03_02] [Medline: 9358596]
12. Mahone IH, Farrell S, Hinton I, Johnson R, Moody D, Rifkin K, Region Ten CSB/UVA School of Nursing Partnership. Shared decision making in mental health treatment: qualitative findings from stakeholder focus groups. Arch Psychiatr Nurs 2011 Dec;25(6):e27-e36. [doi: 10.1016/j.apnu.2011.04.003] [Medline: 22114804]
13. Frank AF, Gunderson JG. The role of the therapeutic alliance in the treatment of schizophrenia. Relationship to course and outcome. Arch Gen Psychiatry 1990 Mar;47(3):228-236. [Medline: 1968329]
14. PHAMOUS. Pharmacotherapy monitoring and outcome survey URL: http://www.phamous.eu/home.html [accessed 2011-12-08] [WebCite Cache ID 63mSD7VGc]

15. Makkink S, Kits L. Herstellen doe je zelf. In: Van Hees S, Van der Vlist P, Mulder N, editors. ROM. Van weten naar meten, ROM in de GGZ. Amsterdam: Boom; 2011:97-108.
16. Schaefer B, Nijssen Y, Van Weeghel J. Van meten naar oplossingsgericht werken. In: Van Hees S, Van der Vlist P, Mulder N, editors. ROM. Van weten naar meten, ROM in de GGZ. Amsterdam: Boom; 2011:89-96.
17. Schrank B, Sibitz I, Unger A, Amering M. How patients with schizophrenia use the internet: qualitative study. J Med Internet Res 2010 Dec;12(5):e70 [FREE Full text] [doi: 10.2196/jmir.1550] [Medline: 21169176]
18. Kuosmanen L, Jakobsson T, Hyttinen J, Koivunen M, Välimäki M. Usability evaluation of a web-based patient information system for individuals with severe mental health problems. J Adv Nurs 2010 Dec;66(12):2701-2710. [doi: 10.1111/j.1365-2648.2010.05411.x] [Medline: 20722811]
19. Jones RB, Atkinson JM, Coia DA, Paterson L, Morton AR, McKenna K, et al. Randomised trial of personalised computer based information for patients with schizophrenia. BMJ 2001 Apr 7;322(7290):835-840. [Medline: 11290639]
20. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/appi.ps.61.11.1099] [Medline: 21041348]
21. Bickmore TW, Puskar K, Schlenk EA, Pfeifer LM, Sereika SM. Maintaining reality: relational agents for antipsychotic medication adherence. Interacting with Computers 2010;22(4):276-288. [doi: 10.1016/j.intcom.2010.02.001]
22. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. Br J Psychiatry 1998 Jan;172:11-18. [Medline: 9534825]
23. Priebe S, Huxley P, Knight S, Evans S. Application and results of the Manchester Short Assessment of Quality of Life (MANSA). Int J Soc Psychiatry 1999;45(1):7-12. [Medline: 10443245]
24. Emerencia AC, Krieke van der L, Petkov N, Aiello M. Assessing schizophrenia with an interoperable architecture. 2011 Oct 28 Presented at: Assessing schizophrenia with an interoperable architecture. 20th ACM International Conference on Information and Knowledge Management MIXHS'11; 2011 Oct 28; Glasgow, Scotland, UK. p.82. In press.
25. van der Krieke L, Emerencia AC, Sytema S. An online portal on outcomes for Dutch service users. Psychiatr Serv 2011 Jul;62(7):803. [doi: 10.1176/appi.ps.62.7.803] [Medline: 21724797]
26. Rotondi AJ, Sinkule J, Haas GL, Spring MB, Litsche CM, Newhill CE, et al. Designing websites for persons with cognitive deficits: design and usability of a psychoeducational intervention for persons with severe mental illness. Psychological Services 2007;4(3):202. [doi: 10.1037/1541-1559.4.3.202]
27. Cooper A, Reimann R. Implementation models and mental models. In: Cooper A, Reimann R, editors. About Face 2.0: the essentials of user interface design. Indianapolis: Wiley Publishing; 2003:21-32.
28. Valimaki M, Anttila M, Hatonen H, Koivunen M, Jakobsson T, Pitkanen A, et al. Design and development process of patient-centered computer-based support system for patients with schizophrenia spectrum psychosis. Inform Health Soc Care 2008 Jun;33(2):113-123. [doi: 10.1080/17538150802127207] [Medline: 18604755]
29. Stephanidis C. editor. User interfaces for all. Concepts, methods, and tools. Mahwah: Lawrence Erlbaum Associates; 2001.
30. Hennink M, Hutter I, Bailey A. Participant recruitment. In: Hennink M, Hutter I, Bailey A, editors. Qualitative Research Methods. London: Sage Publications Ltd; 2011:81-107.
31. Nielsen J. What is usability? In: Usability engineering. San Francisco, Calif.: Morgan Kaufmann Publishers; 1993:23-48.
32. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT'93 and CHI'93 Conference on human factors in computing systems. 1993 Presented at: INTERACT'93 and CHI'93 conference; 24-29 April 1993; Amsterdam, The Netherlands;. [doi: 10.1145/169059.169166]
33. Nielsen J. Heuristic evaluation. In: Nielsen J, Mack RL, editors. Usability inspection methods. New York: Wiley; 1994:25.
34. Hennink M, Hutter I, Bailey A. Observation. In: Qualitative Research Methods. London: Sage Publications Ltd; 2011:169-200.
35. 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] [Medline: 20542857]
36. Internet World Stats. Top 58 countries with highest penetration rates URL: http://www.internetworldstats.com/stats9.htm [accessed 2011-10-20] [WebCite Cache ID 62ZgZpqAm]

Abbreviations

ICT: Information and Communication Technology
ROM: Routine Outcome Monitoring
WEGWEIS: name of the web-based support system described in this study. WEGWEIS is a Dutch abbreviation that stands for web environment for empowerment and individual advice.

Edited by B Bewick; submitted 30.08.11; peer-reviewed by L Yardley, J Freedman, D Gammon; comments to author 27.09.11; revised version received 16.12.11; accepted 19.12.11; published 06.02.12.

Please cite as:
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;14(1):e24
URL: http://www.jmir.org/2012/1/e24/
doi:10.2196/jmir.1921
PMID:22311883

Original Paper

Interpreting the Outcomes of Automated Internet-Based Randomized Trials: Example of an International Smoking Cessation Study

Yan Leykin[1], PhD; Adrian Aguilera[2], PhD; Leandro D Torres[1,3], PhD; Eliseo J Pérez-Stable[4,5], MD; Ricardo F Muñoz[1,3], PhD

[1]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States

[2]School of Social Welfare, University of California, Berkeley, Berkeley, CA, United States

[3]San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, United States

[4]Division of General Internal Medicine, Department of Medicine, University of California, San Francisco, San Francisco, CA, United States

[5]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Ricardo F Muñoz, PhD
San Francisco General Hospital
University of California, San Francisco
1001 Potrero Avenue
Suite 7M
San Francisco, CA, 94110
United States
Phone: 1 415 206 5214
Fax: 1 415 206 8942
Email: ricardo.munoz@ucsf.edu

Abstract

Background: Smoking is one of the largest contributors to the global burden of disease. Internet interventions have been shown to reduce smoking rates successfully. However, improved methods of evaluating effectiveness need to be developed for large-scale Internet intervention trials.

Objective: To illustrate a method to interpret outcomes of large-scale, fully automated, worldwide Internet intervention trials.

Methods: A fully automated, international, Internet-based smoking cessation randomized controlled trial was conducted in Spanish and English, with 16,430 smokers from 165 countries. The randomized controlled trial replicated a published efficacy trial in which, to reduce follow-up attrition, 1000 smokers were followed up by phone if they did not provide online follow-up data.

Results: The 7-day self-reported abstinence rates ranged from 36.18% (2239/6189) at 1 month to 41.34% (1361/3292) at 12 months based on observed data. Given high rates of attrition in this fully automated trial, when participants unreachable at follow-up were presumed to be smoking, the abstinence rates ranged from 13.63% (2239/16.430) at 1 month to 8.28% (1361/16,430) at 12 months. We address the problem of interpreting results with high follow-up attrition rates and propose a solution based on a smaller study with intensive phone follow-up.

Conclusions: Internet-based smoking cessation interventions can help large numbers of smokers quit. Large-scale international outcome studies can be successfully implemented using automated Internet sites. Interpretation of the studies' results can be aided by extrapolating from results obtained from subsamples that are followed up by phone or similar cohort maintenance methods.

Trial Registration: ClinicalTrials.gov NCT00721786; http://clinicaltrials.gov/ct2/show/NCT00721786 (Archived by WebCite at http://www.webcitation.org/63mhoXYPw)

(J Med Internet Res 2012;14(1):e5) doi:10.2196/jmir.1829

KEYWORDS

Smoking cessation; tobacco use; Internet intervention; evidence-based intervention; attrition; effectiveness trial

Introduction

One billion tobacco-related deaths are projected for the 21st century, 80% of which will occur in low- and middle-income countries [1]. Current health care approaches for smoking cessation include nicotine replacement therapy (eg, nicotine patch or gum; quit rate: 14%–24%), other prescription medication (eg, bupropion or varenicline; quit rates: up to 25% and 30%, respectively), and psychosocial interventions (quit rates: 15%–27%) [2-4]. However, these are expensive, unavailable to many who need them, and consumable in terms of both the actual product and the time of clinicians providing treatment. Highly scalable, widely accessible, nonconsumable, evidence-based smoking cessation methods must be developed and evaluated to reduce smoking rates. Large-scale automated self-help Internet interventions are one such method; however, the implementation of these trials presents several practical and methodological challenges. This report presents a promising approach to the implementation of automated, worldwide Internet randomized trials and the interpretation of their results, with attention to assessing effectiveness in addition to efficacy.

Limitations of Current Smoking Cessation Approaches

Internet-based interventions have several advantages, including time and cost effectiveness, almost unlimited scalability, increased intervention fidelity, ease of updates and expansions to conform to the most up-to-date research, and the ability to make them available across the world. With such penetration, even small improvements in likelihoods of quitting smoking can have profound effects on public health relative to the cost of the intervention (see the RE-AIM framework [5-11]). Reports indicate that, in fact, Internet interventions can obtain quit rates comparable with those of other treatment modalities [12-19].

A key benefit of Internet interventions is the automation of delivery. However, with few exceptions [13,18], most trials rely on live personnel to conduct follow-up assessments via phone or email contacts with participants. Insofar as interaction with live personnel may affect outcomes, these trials depart from the fully automated framework. The two main reasons for sacrificing the benefits of automation and fidelity of administration, while incurring considerable costs to conduct live follow-up, are attrition and the analytical convention of imputing smoking status to missing data when presenting cessation trial outcomes.

The progression of Internet intervention studies to address a particular health problem at a worldwide level generally begins with face-to-face clinical trials. The interventions developed at this stage are then adapted for delivery via the Web. Online randomized controlled trials with strong cohort maintenance efforts, such as using staff to send personalized email or to make phone calls to reach participants who do not respond to automated follow-up assessments, can provide estimates of outcome that approximate traditional methods. We suggest that the next step ought to be very-large-scale randomized trials, conducted in a fully automated fashion, to reflect as closely as possible the routine dissemination of Internet interventions that can be made available to anyone in the world, with minimal staffing. However, such large-scale trials generally cannot afford individual live follow-up. This report presents a method that may help researchers in the field to estimate effectiveness data of self-help automated interventions.

Attrition

The motivation needed to enter traditional face-to-face trials is high: people either actively seek them out, respond to an advertisement by calling and visiting a clinic, or are directly recruited from preexisting registries based on demographic, behavioral, or clinical factors. In contrast, those signing up for a Web-based trial generally do so via a Web search and clicking on a link. Of the thousands who visit the website, few will elect to join, fewer still will make adequate use of the intervention, and only a minority will respond to automated follow-up invitations. The difference in effort involved to enter an Internet trial versus a face-to-face trial makes comparisons between the two problematic. Website visitors are more akin to persons reading an advertisement for a trial, most of whom will not actually call or visit the study clinic. Those filling out an online eligibility questionnaire are similar to those calling a phone number to inquire more about a traditional outcome study. Signing up for an online trial takes little effort; although many online participants are likely curious about the Internet trial, they may not be as committed to participating as those signing up after traveling to a study clinic. Once people enter into the study, it is extremely easy for them to drop out of an Internet trial, since there has been no direct personal contact with study staff. Researchers in the field need to reconsider how best to interpret findings that involve large attrition to systematically study the effectiveness of Internet interventions as they would be routinely used in practice, rather than as part of a well-staffed randomized controlled trial.

Attrition is a recognized concern in Internet trials [20], which affects interpretation of results in two ways: (1) if most participants drop out, the remaining sample is highly self-selected and may not be representative of the original visitors, and (2) if participants do not complete the intervention, but respond to follow-up, the outcome data won't represent the intervention's potential. Of course, there are also parallels in face-to-face trials: participants are also highly self-selected and not a representative sample of all who have the disorder being treated.

The "Missing = Smoking" Convention

The usual strategy for determining quit rates in a cessation trial is the *missing = smoking* (M=S) convention, a variant of the intent-to-treat analysis, which presumes that all participants unreachable at follow-up are smoking. This is similar to the last observation carried forward (LOCF) convention; however, the M=S convention is more conservative, for two reasons. First, the LOCF convention permits *responded-to-treatment* observations to be carried forward as well as *nonresponse*, whereas M=S presumes that every dropout is a treatment failure. Second, because cessation trial outcomes are dichotomous (quit vs did not quit), the degree of response/nonresponse (eg, fewer cigarettes) cannot be captured by M=S.

The outcomes of cessation trials therefore largely depend on the completeness of follow-up data. For example, suppose the *true* quit rate for a hypothetical intervention is 20%. Three trials

assessing the effect of that intervention with follow-up rates of 100%, 70%, and 40% would yield M=S outcomes of 20%, 14%, and 8%, respectively, prompting widely differing conclusions about efficacy. Since 14%–22% can be expected with a nicotine patch [4] and 4%–8% can be obtained with a placebo patch, the possible M=S outcome implications vary significantly. Because automated Internet trials have inherently high dropout rates, the M=S convention may be more reflective of follow-up success than of treatment efficacy.

An Illustration of an Interpretable Internet-Based Cessation Trial

With live follow-up (eg, phone calls), it is possible for geographically limited Internet trials to obtain follow-up rates of up to 78%, which is comparable with face-to-face trials [21]. However, live follow-ups for large-scale, worldwide Internet trials are costly and logistically difficult. Conversely, allowing the logistical limitations of live follow-up procedures to constrain the number of participants compromises scope and reach, limiting the public health applications of an Internet trial.

We propose one possible model of structuring an Internet trial that may help assess effectiveness of a trial once efficacy is established. In 2009, Muñoz and colleagues reported on the outcome of a Web-based smoking cessation trial conducted in Spanish and English (n = 1000) [14]. Live follow-up was used with those who did not provide data after an automated reminder, obtaining follow-up rates of 68% at 12-month follow-up. At 12 months, 20% of Spanish- and 21% of English-speaking participants were no longer smoking (M=S). After random allocation of the first 1000 participants, live follow-up ended, but the rest of the online intervention study was left exactly the same, with the goal of conducting a larger trial to demonstrate the demand for and the reach of the intervention as delivered in a fully automated format. Here, we report on the results of the fully automated portion of that trial.

Conducting a smaller and logistically feasible live follow-up trial followed by or concurrently with a larger fully automated trial can address the concerns of cost versus scope mentioned previously. The goal of the current study was therefore to illustrate the use of this approach in interpreting the outcomes of a fully automated trial. We used the outcomes of the Muñoz et al [14] trial to interpret the results of the current fully automated trial, and we tested three hypotheses to determine whether a more complex intervention would increase quit rates.

Methods

Participants

Recruitment procedures were the same as described elsewhere [14]. Google AdWords ads were the main means for recruiting participants. Eligible participants were 18+ years of age, smoking 5+ cigarettes per day, with regular (1+ times/week) access to email and Internet, intending to quit in the next month. Of the visitors screened for eligibility, 16,475/78,623 (20.95%) were ineligible, 1052/78,623 (1.34%) were <18 years old, 2738/78,623 (3.48%) smoked <5 cigarettes/day, 9875/78,623 (12.56%) were not ready to quit, and 4646/78,623 (5.91%) had no email address. Participants were not paid for their participation in the study. Participants were recruited from November 2005 to September 2009.

Study Procedures

Study procedures are described in detail elsewhere [14]. Briefly, visitors to the site completed brief demographics and eligibility questionnaires. Eligible participants viewed and e-signed a consent document, which detailed the study procedures, including randomization. Consenting participants completed baseline questionnaires. To select out one-time visitors, participants were asked to return 3 times over the next 7 days and report the number of cigarettes smoked. Those meeting this requirement set their quit dates, were automatically randomly assigned to 1 of 4 conditions, and were given access to the website. Participants were sent automated follow-up assessment emails at 1, 3, 6, and 12 months after their quit date. Only self-reported smoking data were gathered, for three reasons: (1) biochemical verification was not feasible for a very large worldwide trial, (2) the fully automated nature of the trial precluded additional participant contact, and (3) there is growing evidence that self-report is sufficient for nonintensive interventions [22-24].

The only difference from the procedures described in the 2009 [14] paper, wherein research assistants either called or emailed those who did not submit follow-up data after the automated email contact, was that in the present study we did not use live follow-up.

Study Conditions

As in the 2009 trial [14], participants were randomly assigned to 1 of 4 conditions. The website used a preprogrammed algorithm for random assignment using stratified randomization, with gender and history of major depressive episodes (MDEs) as stratification variables. Condition 1 contained the most basic elements, and conditions 2–4 incrementally added further elements.

The 4 arms (conditions) of the trial were the following:

1. A noninteractive, static smoking cessation guide (Guía para dejar de fumar [25-27]), a cigarette counter, and an online journal.

2. Condition 1, plus individually timed email messages: preprogrammed emails with links to sections of smoking cessation guide timed to quit date [28].

3. Condition 2, plus an 8-session cognitive–behavioral mood management course (based on Lewinsohn et al [29]).

4. Condition 3, plus a virtual participant-driven, unmoderated support group (an asynchronous bulletin board).

Hypotheses

We retained three specific hypotheses regarding the outcome of the intervention from the 2009 [14] study and tested them in the fully automated sample: (1) conditions 2, 3, and 4 will outperform condition 1, (2) condition 4 will obtain the best quit rates, followed by condition 3, followed by condition 2, followed by condition 1, and (3) conditions 3 and 4 (containing mood management) will outperform conditions 1 and 2.

Measures

A *demographic questionnaire* included age, gender, race/ethnicity, education, income, and marital status.

A *smoking questionnaire* included age when the participant started smoking, age when smoking regularly, number of cigarettes per day, confidence in quitting, and smoking exposure.

The *Fagerström Test for Nicotine Dependence* (FTND) [30] is a commonly used 6-item test of nicotine dependence, with a range from 0 to 10.

The *MDE Screener* (Mood Screener) [31] screens for the presence of the 9 symptoms of current and past MDEs according to the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, as well as for criterion C (significant impairment in functioning). This instrument has been shown to have good agreement with the PRIME-MD and with clinician-administered interviews [32-34].

The *Center for Epidemiologic Studies Depression* scale (CES-D) [35] is a 20-item self-report scale designed to measure the current level of depressive symptoms.

Statistical Analyses

To test hypotheses 1 (condition 1 will result in worse outcomes) and 3 (conditions with mood management will result in better outcomes), we conducted repeated binary logistic regressions. The quit rates were predicted from the intervention condition assignment (1 versus others for hypothesis 1; 1 and 2 versus 3 and 4 for hypothesis 3), covarying participant demographic characteristics (gender, age, education, and race), language (English or Spanish), depression (CES-D score and presence of current or past MDE), and level of addiction (FTND). We conducted these analyses twice: once with the M=S assumption, and the other with observed data (without the M=S assumption).

To test hypothesis 2—that intervention conditions would yield incrementally better outcomes—we constructed binary logistic regression models, predicting the 7-day quit rate at 1, 3, 6, and 12 months. The model predictors were the same as those used for repeated measures analyses, described above. As above, these analyses were conducted twice: once with the M=S assumption, and the other with observed data (without the M=S assumption).

Due to the considerable size of the sample (n = 16,430), we elected to report significance only if we obtained a *P* value less than .01, to reduce type I error.

Results

Sample Characteristics

Participants were 16,430 smokers (3332 English- and 13,098 Spanish-speaking), aged 18 to 84 (mean 36.2, SD 10.7), from 165 countries. The three most-represented countries for English speakers were the United States (n = 1251), India (n = 358), and South Africa (n = 306). The three most-represented Spanish-speaking countries were Spain (n = 4341), Argentina (n = 2513), and Mexico (n = 2100). Just over half of participants were men (8638/16,349, 52.84%), and most were well educated (12,628/16,379, 77.10% with at least some college education), gainfully employed (12,960/16,415, 78.95% at least part-time), and married or living as married (8846/16,403, 53.93%).

Participants reported having smoked for 20.6 years, on average (SD 10.9), smoking on average 1 pack per day (mean 19.6, SD 9.9 cigarettes). The average age at first cigarette was 15.6 (SD 3.2) years, and the average age of smoking regularly (first 5 packs) was 18.6 (SD 4.3). The average level of nicotine dependence, as measured by the FTND, was 5.2 (SD 2.5), indicating moderate dependence, and similar to face-to-face smoking cessation trials [36-39].

Participant characteristics for each condition are shown in Table 1.

Table 1. Participant characteristics, by condition[a]

	Condition 1 (n = 4118)	Condition 2 (n = 4097)	Condition 3 (n = 4110)	Condition 4 (n = 4105)	*P* value[b]
Male, n (%)	2168/4102 (52.85%)	2150/4080 (52.70%)	2165/4088 (52.96%)	2155/4079 (52.83%)	1.00
Age (years), mean (SD)	36.1 (11.4)	36.3 (11.8)	36.5 (14.5)	36.4 (13.6)	.47
Some college or more, n (%)	3167/4107 (77.11%)	3141/4088 (76.83%)	3173/4091 (77.56%)	3147/4093 (76.89%)	.93
White, n (%)	2803/4086 (68.60%)	2801/4069 (68.84%)	2796/4076 (68.60%)	2802/4075 (68.76%)	.37
Spanish-speaking, n(%)	3284/4118 (79.75%)	3263/4097 (79.64%)	3275/4110 (79.68%)	3276/4105 (79.81%)	1.00
Employed, n (%)	3217/4115 (78.18%)	3237/4091 (79.12%)	3268/4109 (79.53%)	3238/4100 (78.98%)	.36
Married or partnered, n (%)	2206/4112 (53.65%)	2209/4089 (54.02%)	2234/4103 (54.45%)	2197/4099 (53.60%)	.86
CES-D[c] score, mean (SD)	16.9 (12.1)	17.0 (12.0)	16.8 (12.4)	16.9 (12.4)	.74
Current or past major depressive episode, n (%)	1276/4109 (31.05%)	1275/4091 (31.17%)	1280/4103 (31.20%)	1277/4101 (31.14%)	1.00
Cigarettes/day, mean (SD)	19.4 (9.9)	19.8 (10.2)	19.5 (10.1)	19.6 (9.7)	.36
Age started smoking (years), mean (SD)	15.5 (3.2)	15.6 (3.2)	15.5 (3.2)	15.6 (3.4)	.46
Age regular smoker (years), mean (SD)	18.6 (4.4)	18.5 (4.0)	18.6 (4.3)	18.7 (4.4)	.19
Years smoked, mean (SD)	20.5 (10.8)	20.6 (10.9)	20.7 (10.9)	20.6 (11.0)	.81
FTND[d] score, mean (SD)	5.2 (2.5)	5.3 (2.5)	5.2 (2.5)	5.2 (2.5)	.41

[a] Conditions were as follows: condition 1: a noninteractive smoking cessation guide, cigarette counter, and an online journal; condition 2: condition 1, plus individually timed email messages; condition 3: condition 2, plus an 8-session cognitive–behavioral mood management course; and condition 4: condition 3, plus a virtual participant-driven support group.

[b] *P* values were determined via 1-way analyses of variance for continuous variables, and via Pearson chi-squares for categorical variables.

[c] Center for Epidemiologic Studies Depression scale.

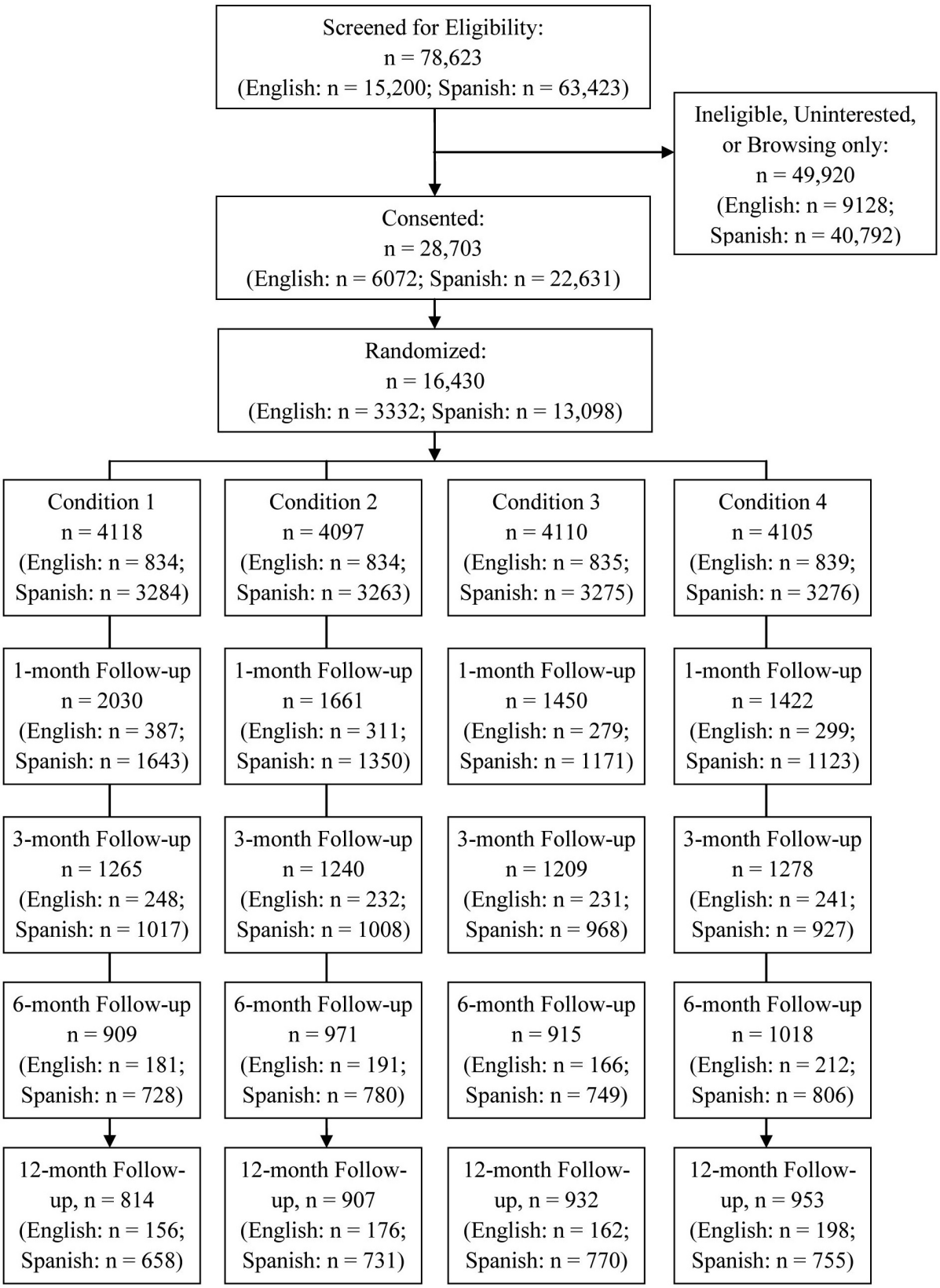
[d] Fagerström Test for Nicotine Dependence.

Attrition

The progression of participants through the study is outlined in Figure 1. Of the over 150,000 participants who visited our website, 78,623 provided enough data to evaluate their eligibility, 28,703 signed consent, and 16,430 completed baseline assessments and the washout period and underwent random assignment.

The current study relied solely on automated emailed reminders to obtain follow-up data. For month 1 follow-up, 6563/16,430 (40.0%) participants provided data. This number was reduced to 4992/16,430 (30.38%), 3813/16,430 (23.21%), and 3606/16,430 (21.95%) for follow-ups at months 3, 6, and 12, respectively. These numbers were comparable with those obtained in the earlier [14] study, where 38%, 30%, 27%, and 23% of participants who never received any live follow-up returned at months 1, 3, 6, and 12, respectively.

Figure 1. CONSORT diagram for progression of participants through the fully automated Internet stop smoking trial.



Abstinence Rates

Based on observed data, 1 month after enrollment, 36.18% (2239/6189) reported not having smoked in the past 7 days (Table 2). A similar proportion of participants reported 7-day abstinence at months 3 (1797/4566, 39.36%), 6 (1478/3508, 42.13%), and 12 (1361/3292, 41.34%). A somewhat smaller proportion of participants reported not having smoked in the past 30 days (1640/6182, 26.53%; 1465/4562, 32.11%; 1243/3504, 35.47%; and 1211/3286, 36.85% at 1, 3, 6, and 12

months, respectively). Using the M=S convention, the cessation rates observed in this study were modest. The 7-day abstinence rates at 1, 3, 6, and 12 months, respectively, were 13.63% (2239/16,430), 10.94% (1797/16,430), 9.00% (1478/16,430), and 8.28% (1361/16,430). The respective 30-day abstinence rates were 9.98% (1640/16,430), 8.92% (1465/16,430), 7.57% (1243/16,430), and 7.37% (1211/16,430). In the 45 months of this intervention being available online, 3,489 individuals reported having quit for at least 7 days (about 18 per week over the course of the study), and 2,786 individuals (about 14 per week) reported having quit for at least 30 days.

Table 2. Overall self-reported abstinence rates (% quit) in an online sample of 16,430 consented smokers

	1-month follow-up		3-month follow-up		6-month follow-up		12-month follow-up	
Completed follow-ups, n (%)	6563/16,430 (39.95%)		4992/16,430 (30.38%)		3813/16,430 (23.21%)		3606/16,430 (21.95%)	
	7 days	30 days	7 days	30 days	7 days	30 days	7 days	30 days
Observed								
n	2239/6189	1640/6182	1797/4566	1465/4562	1478/3508	1243/3504	1361/3292	1211/3286
%	36.18%	26.53%	39.36%	32.11%	42.13%	35.47%	41.34%	36.85%
M=S[a]								
n	2239/16,430	1640/16,430	1797/16,430	1465/16,430	1478/16,430	1243/16,430	1361/16,430	1211/16,430
%	13.63%	9.98%	10.94%	8.92%	9.00%	7.57%	8.28%	7.37%

[a] Missing observations are presumed to be smoking.

Intervention Conditions

We noted several differences between treatment conditions (Table 3). With observed data (ie, without the M=S assumption), at 1-month follow-up, condition 1 performed significantly poorer than all other conditions, in partial support for hypothesis 1 (Wald $\chi^2_3 = 80.7$, $P < .001$). No significant differences between conditions were noted at months 3 (Wald $\chi^2_3 = 10.3$, $P = .02$), 6 (Wald $\chi^2_3 = 5.8$, $P = .12$), and 12 (Wald $\chi^2_3 = 7.5$, $P = .06$). With M=S analyses, significant differences were observed at months 6 (Wald $\chi^2_3 = 14.8$, $P = .002$) and 12 (Wald $\chi^2_3 = 13.0$, $P = .005$), such that conditions 2 and 4 outperformed conditions 1 and 3.

Observing the quit rates, it is clear that hypothesis 2—that conditions would result in incremental improvements in quit rates—is not supported. To test the two other hypotheses, we conducted repeated-measures logistic regressions, with the same covariates as in the simple logistic regressions above. Hypothesis 1 was largely supported. With observed data, condition 1 resulted in lower quit rates than conditions 2, 3, and 4 (Wald $\chi^2_1 = 30.1$, $P < .001$, beta = –.28, 95% confidence interval [CI], –.39 to –.18). A similar result was observed with M=S data, though the result did not cross the significance level set for this study (Wald $\chi^2_1 = 6.1$, $P = .01$, beta = –.12, 95% CI –.21 to –.02). Hypothesis 3—that mood management conditions would result in higher quit rates—was supported only with observed data (Wald $\chi^2_1 = 9.5$, $P = .002$, beta = –.14, 95% CI –.23 to –.05), but not with the M=S data (Wald $\chi^2_1 = .0$, $P = .96$, beta = .00, 95% CI –.08 to .78).

Table 3. 7-day quit rates (n, %) by intervention condition

	Condition[a]				*P* value
	1 (cessation guide)	2 (1 + email messages)	3 (2 + mood management)	4 (3 + virtual group)	
Observed					
Month 1	526/1912 (27.51%)	611/1578 (38.72%)	550/1371 (40.12%)	552/1328 (41.57%)	<.001[b]
Month 3	427/1175 (36.34%)	467/1156 (40.40%)	428/1132 (37.81%)	475/1103 (43.06%)	.02
Month 6	327/827 (39.54%)	395/893 (44.23%)	342/845 (40.47%)	414/943 (43.90%)	.12
Month 12	306/730 (41.92%)	355/833 (42.62%)	314/845 (37.16%)	386/884 (43.67%)	.06
Missing = smoking					
Month 1	526/4118 (12.77%)	611/4097 (14.91%)	550/4110 (13.38%)	552/4105 (13.45%)	.03
Month 3	427/4118 (10.37%)	467/4097 (11.40%)	428/4110 (10.41%)	475/4105 (11.57%)	.20
Month 6	327/4118 (7.94%)	395/4097 (9.64%)	342/4110 (8.32%)	414/4105 (10.09%)	.002[b]
Month 12	306/4118 (7.43%)	355/4097 (8.66%)	314/4110 (7.64%)	386/4105 (9.40%)	.005[b]

[a] Conditions were as follows: 1: a noninteractive smoking cessation guide, cigarette counter, and an online journal; 2: condition 1, plus individually timed email messages; 3: condition 2, plus an 8-session cognitive–behavioral mood management course; and condition 4: condition 3, plus a virtual participant-driven support group.

[b] Significant, controlling for demographic characteristics (gender, age, education, race), language of the intervention (English or Spanish), level of addiction (Fagerström Test for Nicotine Dependence [FTND] score), and depression (Center for Epidemiologic Studies Depression scale [CES-D] score and presence of current or past major depressive episodes).

Putting the Outcomes of a Fully Automated Trial in Perspective

For the current trial, the quit rate at 12 months was 8%, assuming M=S. However, the *true* quit rate is unknown: 8% is clearly an underestimate, because the M=S is only an assumption and is highly conservative. As a thought experiment, if the opposite assumption is made—that all those with missing data have quit (M=Quit)—then the quit rate would be 88% (8% observed quit + 80% missing). The M=S is an underestimate of the true quit rate, but M=Quit is clearly an overestimate. Though we can say with confidence that the true quit rate resides between 8% and 88%, this is not very informative (Figure 2).

We can approximate the true quit rate by using the rates from the earlier [14] trial. Because the 2009 trial used exactly the same intervention (plus live follow-up), similar quit rates can be expected. To determine whether the two cohorts are similar, we used binary logistic regression with the same covariates as in the previous models to compare the two most similar subgroups in the two trials: those in the 2009 study who completed all four follow-ups after only an automated email reminder (and were thus never exposed to live follow-up), and those in the current study who also completed all four follow-ups. For these subgroups, the 12-month quit rates were not significantly different (46/96, 48% in the 2009 study vs 714/1326, 54% in the present study; Wald $\chi^2_1 = 1.6$, $P = .21$, odds ratio = 0.75, 95% CI 0.48–1.17), which provides additional credence for the ability to extrapolate results of the current study in the context of the 2009 trial. Only two covariates crossed the significant threshold. One was depression history (Wald $\chi^2_2 = 13.8$, $P = .001$), with participants without a history of depression appearing to quit at higher rates (603/1043, 57.8%) than those with past (92/213, 43.2%) or current (62/156, 40%) depression. The other was FTND score, with those scoring higher being less likely to quit (Wald $\chi^2_1 = 7.0$, $P = .01$, odds ratio = 0.94, 95% CI 0.90–0.99).

The 2009 [14] study obtained an average quit rate of 21% at 12 months assuming M=S, or, assuming M=Quit, 52%. The true rate for the current study therefore most likely resides between 21% and 52%, which is considerably more informative than the 8%–88% interval (Figure 2).

The interval can be narrowed down further. The observed quit rate for the current study at 12 months is 41%; in the 2009 [14], it was 30%. However, that 30% was based on both automated and live follow-up responders, and the reported quit rates of automated responders were about 70% higher than that of live responders (16.3% vs 9.8%, respectively, across all follow-ups). The current study's observed quit rate (41%) is therefore an overestimate of the true quit rate, as everyone who provided data did so with automated follow-up. The most likely conclusion about the true quit rate in the current study is that the upper bound is 30% (observed quit rate in the 2009 study) and the lower bound is 21%, as illustrated in Figure 2.

Figure 2. Most likely quit rate range extrapolated from the current trial and an identical trial with live follow-up (Muñoz et al [14]). M=Quit: missing observations are presumed quit ([reported quit + missing]/all assigned to condition); M=S: missing observations are presumed smoking (reported quit/all assigned to condition); Observed: missing observations are excluded (reported quit/[reported quit + reported smoking).

Discussion

In this paper we have highlighted the problems of attrition in international Internet trials, especially in the context of the M=S convention, and offered a way to reconcile the demands of needing to employ costly means of follow-up with the advantages that the breadth of a very-large-scale automated trial allows. By referencing the identically conducted trial, with the only difference being live follow-up for those who did not respond to automated email reminders, we estimated the true quit rate for the current trial to lie between 21% and 30% of participants. We also found that more complex versions of the intervention resulted in better cessation rates than a static online smoking cessation guide, suggesting that some level of complexity and personalization may be helpful in Internet interventions.

Internet interventions are a relatively new form of health-promoting behavior change interventions that are likely to grow considerably due to the benefits of reach and cost effectiveness. To ensure that these interventions are improving health outcomes, they must be tested to ensure a strong evidence base. Indeed, Internet-based interventions are increasingly evidence based [40,41]. These interventions will reach a larger proportion of the world with increased Internet penetration [42] and improving technology such as integration of mobile devices and linkages with electronic health records. However, novel interventions call for novel methods of their evaluation, especially where conventional methods fall short.

One of the most significant benefits of Internet interventions is their cost effectiveness due to sustainability and nonconsumability. The cost of creating an effective Internet site for smoking cessation is relatively modest, about US $55,000 in our case. Conducting the randomized trials to evaluate its effectiveness costs much more. Once the original efficacy trial with live follow-up was completed, however, leaving the site open to conduct the fully automated randomized trial reported here was relatively inexpensive. We estimate that it cost about US $120 per participant who reported quitting successfully by at least one follow-up point. If the nicotine patch had been used, assuming a 20% quit rate with the patch, at a cost of US $3.91 per day [4] for 10 weeks, the cost per participant who quit successfully using the patch would have been about US $1370.

The benefit of Internet intervention trials may be undervalued if methods for their evaluations underestimate their effects. Though sophisticated statistical procedures that may tackle missing data exist (eg, multiple imputation), they may not be accurate when the proportion of missing data is very large, as is often the case in fully automated Internet studies. Therefore, we have proposed a hybrid solution that includes conducting a smaller trial with aggressive follow-up using live methods (eg, phone calls) to assess for an intervention's efficacy followed by a larger naturalistic trial to assess effectiveness. This method would allow for the rigorous testing of efficacy by ensuring high follow-up rates with a smaller trial. The larger trial would then allow one to take advantage of the wide reach and automated nature of Internet interventions to assess for the

overall impact of the study by extrapolating results from the smaller trial onto a larger sample. Thus, the 2009 [14] study first established the efficacy of the intervention, and this paper has outlined the second step of our proposed method by highlighting effectiveness of the same intervention delivered in an automated format. These methods can also be applied to prevention trials that have similar issues in that they use dichotomous outcomes and similar conventions when people drop out of a study or are lost to follow-up. Applying a method that first assesses efficacy and then focuses on broad reach would better inform the potential impact of large-scale public health campaigns that are difficult to interpret due to difficulties with and cost of follow-up assessments.

There are limitations to our study and the way we have used our proposed 2-step method. In both studies, smoking was assessed via self-report rather than biomedical validation measures; however, this is the recommended approach in large-scale community trials [24]. Participants in the efficacy trial [14] were the first 500 Spanish speakers and 500 English speakers randomly assigned to the 4 study conditions. History effects were therefore not controlled, though the two cohorts were found to be comparable. In future studies, the follow-up cohort should be selected randomly across time from the large sample of participants in automated self-help studies. Lastly, the actual outcomes of participants who did not provide follow-up data are unknown. We have made the case that informed estimates can be made, when they are based on efficacy data in a subsample with rigorous follow-up. In some parts of the world, Internet access is available only to those of higher socioeconomic status. This is rapidly changing, however, with the growth of Internet penetration being the highest in the developing countries [42]. Finally, the majority of participants were non-US Spanish speakers. The results may not generalize to other populations.

Internet-based interventions for health problems are becoming increasingly popular due to their enormous reach and cost effectiveness: no other medium permits conducting a randomized trial of an empirically supported intervention for over 16,000 individuals across 165 countries at such low cost. However, in testing these interventions via randomized controlled trials, particularly when assessing dichotomous outcomes, it is necessary to develop new methods of analysis that are able to fully reflect the true impact and effectiveness of large-scale, international public health Internet interventions.

Future directions involve carrying out outcome studies that are more generalizable to how Internet interventions would be used outside of a strict randomized trial context. Specifically, users of such sites are likely to pick and choose among intervention elements provided by the sites. Thus, the next step after randomized trials ought to be participant preference trials, in which users are provided access to all elements of the interventions that were found to be reasonably effective within a randomization context, and allowed to use the elements they prefer. Our team is conducting such a study, which we believe would best estimate the effectiveness of a self-help automated Internet intervention that would be made available at no charge to anyone in the world who wanted to use it.

Researchers in the Internet intervention field should consider adopting this approach, namely a progression of studies, from strict efficacy randomized trials (with live follow-ups to reduce attrition), to fully automated randomized trials (to approximate how a self-help site would be used), proceeding to participant preference effectiveness studies (in which all elements tested in the earlier randomized trials are made available to all participants). Such an approach would contribute to the use of evidence-based Internet interventions to reduce health disparities worldwide [43].

Acknowledgments

This research was supported by grant 13RT-0050 from the Tobacco-Related Disease Research Program, by an infrastructure grant from the University of California Committee on Latino Research to the University of California/San Francisco General Hospital Latino Mental Health Research Program (LMHRP; Muñoz, Principal Investigator), by grants from the Tobacco Research Network Program, Fogarty International Center, National Institute of Drug Abuse (No.TW05935, Pérez-Stable, Principal Investigator), and from the National Cancer Institute for Redes en Acción (U01CA86117, Pérez-Stable, Principal Investigator), National Institutes of Health, by NIMH training grant to Adrian Aguilera (MH 5T32MH018261-27; Patricia Arean, Principal Investigator), and by NIMH grant 5K08MH091501 (Leykin, Principal Investigator). The authors thank the Center for Health and Community (Nancy Adler, Director) for providing office space and additional resources. Special thanks go to Google, Inc for awarding the UCSF/SFGH LMHRP an AdWords grant (Muñoz, Principal Investigator), which provided us with the ability to recruit smokers worldwide using Google sponsored links.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH V1.6 checklist [44].

[PDF File (Adobe PDF File), 810KB - jmir_v14i1e5_app1.pdf]

References

1. World Health Organization. WHO Report on the Global Tobacco Epidemic 2008: The MPOWER Package. Geneva: World Health Organization; 2008.
2. Fiore MC, Smith SS, Jorenby DE, Baker TB. The effectiveness of the nicotine patch for smoking cessation. A meta-analysis. JAMA 1994;271(24):1940-1947. [Medline: 8201739]
3. Lando HA, McGovern PG, Barrios FX, Etringer BD. Comparative evaluation of American Cancer Society and American Lung Association smoking cessation clinics. Am J Public Health 1990 May;80(5):554-559. [Medline: 2327531]
4. 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] [Medline: 16046655]
5. Glasgow RE. Evaluation of theory-based interventions: the RE-AIM model. In: Glanz K, Rimer BK, Lewis FM, editors. Health Behavior and Health Education: Theory, Research, and Practice. 3rd edition. San Francisco, CA: Jossey-Bass; 2002.
6. Glasgow RE, Klesges LM, Dzewaltowski DA, Bull SS, Estabrooks P. The future of health behavior change research: what is needed to improve translation of research into health promotion practice? Ann Behav Med 2004 Feb;27(1):3-12. [doi: 10.1207/s15324796abm2701_2] [Medline: 14979858]
7. Glasgow RE, Klesges LM, Dzewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. Health Educ Res 2006 Oct;21(5):688-694 [FREE Full text] [doi: 10.1093/her/cyl081] [Medline: 16945984]
8. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. Am J Public Health 2003 Aug;93(8):1261-1267. [Medline: 12893608]
9. Glasgow RE, Nelson CC, Strycker LA, King DK. Using RE-AIM metrics to evaluate diabetes self-management support interventions. Am J Prev Med 2006 Jan;30(1):67-73. [doi: 10.1016/j.amepre.2005.08.037] [Medline: 16414426]
10. Glasgow RE, Orleans CT, Wagner EH. Does the chronic care model serve also as a template for improving prevention? Milbank Q 2001;79(4):579-612, iv. [Medline: 11789118]
11. 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]
12. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]
13. Etter JF. Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial. J Med Internet Res 2005;7(1):e2 [FREE Full text] [doi: 10.2196/jmir.7.1.e2] [Medline: 15829474]
14. 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] [Medline: 19640833]
15. 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] [Medline: 16497602]
16. Pike KJ, Rabius V, McAlister A, Geiger A. American Cancer Society's QuitLink: randomized trial of Internet assistance. Nicotine Tob Res 2007 Mar;9(3):415-420. [doi: 10.1080/14622200701188877] [Medline: 17365773]
17. 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] [Medline: 18984557]
18. 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]
19. Swartz LH, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
20. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
21. Saul JE, Schillo BA, Evered S, Luxenberg MG, Kavanaugh A, Cobb N, et al. Impact of a statewide Internet-based tobacco cessation intervention. J Med Internet Res 2007;9(3):e28 [FREE Full text] [doi: 10.2196/jmir.9.4.e28] [Medline: 17942390]
22. Velicer WF, Prochaska JO, Rossi JS, Snow MG. Assessing outcome in smoking cessation studies. Psychol Bull 1992 Jan;111(1):23-41. [Medline: 1539088]
23. Strecher VJ, Becker MH, Clark NM, Prasada-Rao P. Using patients' descriptions of alcohol consumption, diet, medication compliance, and cigarette smoking: the validity of self-reports in research and practice. J Gen Intern Med 1989;4(2):160-166. [Medline: 2651602]
24. Hughes JR, Keely JP, Niaura RS, Ossip-Klein DJ, Richmond RL, Swan GE. Measures of abstinence in clinical trials: issues and recommendations. Nicotine Tob Res 2003 Feb;5(1):13-25. [Medline: 12745503]
25. 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]
26. National Cancer Institute. Guia Para Dejar de Fumar. NIH Publication No. 02-3001. Washington, DC: US Government Printing Office; 2002.

27. Pérez-Stable EJ, Sabogal F, Marín G, Marín BV, Otero-Sabogal R. Evaluation of "Guia para Dejar de Fumar," a self-help guide in Spanish to quit smoking. Public Health Rep 1991;106(5):564-570. [Medline: 1910191]
28. 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] [Medline: 15064291]
29. Lewinsohn PM, Muñoz RF, Youngren MA, Zeiss AM. Control Your Depression. New York, NY: Simon and Schuster; 1992.
30. 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]
31. Muñoz RF. University of California, San Francisco, School of Medicine, Latino Mental Health Research Program. 2010 Jan 25. The Mood Screener: a Major Depression Episode Screener URL: http://www.medschool.ucsf.edu/latino/manuals.aspx#MoodScreener [accessed 2011-03-23] [WebCite Cache ID 5xQgNI5vK]
32. 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]
33. 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]
34. 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(1):57-64. [doi: 10.1027/1015-5759.24.1.57]
35. 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. [doi: 10.1177/014662167700100306]
36. Hall SM, Humfleet GL, Reus VI, Muñoz RF, Hartz DT, Maude-Griffin R. Psychological intervention and antidepressant treatment in smoking cessation. Arch Gen Psychiatry 2002 Oct;59(10):930-936 [FREE Full text] [Medline: 12365880]
37. Hall SM, Muñoz RF, Reus VI. Cognitive-behavioral intervention increases abstinence rates for depressive-history smokers. J Consult Clin Psychol 1994 Feb;62(1):141-146. [Medline: 8034816]
38. 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]
39. Hall SM, Reus VI, Muñoz RF, Sees KL, Humfleet G, Hartz DT, et al. Nortriptyline and cognitive-behavioral therapy in the treatment of cigarette smoking. Arch Gen Psychiatry 1998 Aug;55(8):683-690 [FREE Full text] [Medline: 9707377]
40. Marks IM, Cavanagh K, Gega L. Hands-on Help: Computer-Aided Psychotherapy. Hove, UK: Psychology Press; 2007.
41. Moser DJ, Reese RL, Hey CT, Schultz SK, Arndt S, Beglinger LJ, et al. Using a brief intervention to improve decisional capacity in schizophrenia research. Schizophr Bull 2006 Jan;32(1):116-120 [FREE Full text] [doi: 10.1093/schbul/sbi066] [Medline: 16177273]
42. Miniwatts Marketing Group. 2010. Internet Usage Statistics: The Internet Big Picture: World Internet Users and Population Stats URL: http://internetworldstats.com/stats.htm [accessed 2011-03-24] [WebCite Cache ID 5xQgrpgBf]
43. 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] [Medline: 21169162]
44. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CES-D: Center for Epidemiologic Studies Depression scale
CI: confidence interval
FTND: Fagerström Test for Nicotine Dependence
LOCF: last observation carried forward
M=Quit: missing observations presumed quit
M=S: missing observations presumed smoking
MDE: major depressive episode

Edited by G Eysenbach; submitted 16.04.11; peer-reviewed by B Carlini; comments to author 02.09.11; revised version received 28.09.11; accepted 18.10.11; published 07.02.12.

Please cite as:
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
URL: http://www.jmir.org/2012/1/e5/
doi:10.2196/jmir.1829
PMID:22314016

Original Paper

Breathe Easier Online: Evaluation of a Randomized Controlled Pilot Trial of an Internet-Based Intervention to Improve Well-being in Children and Adolescents With a Chronic Respiratory Condition

Peter A Newcombe[1], BEd, BA(Hons), PhD; Tamara L Dunn[1], BA(Hons), PhD; Leanne M Casey[2], BA(Hons), MClinPsych, PhD; Jeanie K Sheffield[1], BA(Hons), PhD (Clinical Psychology); Helen Petsky[3], BNursing; Sophie Anderson-James[3], BNursing; Anne B Chang[3,4], MPHTM, MBBS, PhD, FRACP

[1]School of Psychology, University of Queensland, Brisbane, Australia

[2]Griffith Health Institute, Behavioural Basis of Health, Griffith University, Brisbane, Australia

[3]Queensland Children's Respiratory Centre, Queensland Children's Medical Research Institute, Royal Children's Hospital, Brisbane, Australia

[4]Child Health Division, Menzies School of Health Research, Darwin, Australia

Corresponding Author:
Peter A Newcombe, BEd, BA(Hons), PhD
School of Psychology
University of Queensland
Bldg 24, Room s323
Brisbane, 4072
Australia
Phone: 61 7 33469643
Fax: 61 7 33654466
Email: newc@psy.uq.edu.au

Abstract

Background: Chronic respiratory illnesses are the most common group of childhood chronic health conditions and are overrepresented in socially isolated groups.

Objective: To conduct a randomized controlled pilot trial to evaluate the efficacy of Breathe Easier Online (BEO), an Internet-based problem-solving program with minimal facilitator involvement to improve psychosocial well-being in children and adolescents with a chronic respiratory condition.

Methods: We randomly assigned 42 socially isolated children and adolescents (18 males), aged between 10 and 17 years to either a BEO (final n = 19) or a wait-list control (final n = 20) condition. In total, 3 participants (2 from BEO and 1 from control) did not complete the intervention. Psychosocial well-being was operationalized through self-reported scores on depression symptoms and social problem solving. Secondary outcome measures included self-reported attitudes toward their illness and spirometry results. Paper-and-pencil questionnaires were completed at the hospital when participants attended a briefing session at baseline (time 1) and in their homes after the intervention for the BEO group or a matched 9-week time period for the wait-list group (time 2).

Results: The two groups were comparable at baseline across all demographic measures (all $F < 1$). For the primary outcome measures, there were no significant group differences on depression ($P = .17$) or social problem solving ($P = .61$). However, following the online intervention, those in the BEO group reported significantly lower depression ($P = .04$), less impulsive/careless problem solving ($P = .01$), and an improvement in positive attitude toward their illness ($P = .04$) compared with baseline. The wait-list group did not show these differences. Children in the BEO group and their parents rated the online modules very favorably.

Conclusions: Although there were no significant group differences on primary outcome measures, our pilot data provide tentative support for the feasibility (acceptability and user satisfaction) and initial efficacy of an Internet-based intervention for improving well-being in children and adolescents with a chronic respiratory condition.

Trial registration: Australian New Zealand Clinical Trials Registry number: ACTRN12610000214033; http://www.anzctr.org.au/trial_view.aspx?ID=308074 (Archived by WebCite at http://www.webcitation.org/63BL55mXH)

(J Med Internet Res 2012;14(1):e23) doi:10.2196/jmir.1997

KEYWORDS

Internet-based intervention; chronic respiratory condition; psychosocial well-being; children and adolescents; randomized controlled trial

Introduction

Chronic respiratory illnesses include asthma and cystic fibrosis, and are the most common group of childhood chronic conditions [1]. Further, these conditions are more widespread in socially isolated and disadvantaged children and adolescents than they are in higher socioeconomic groups, even after adjusting for traditional risk factors [2,3]. Children and adolescents with a chronic respiratory condition often feel different from their healthy peers due to the necessity of a daily treatment regimen; can have trouble maintaining friendships because of school absences [4]; and can experience psychological difficulties such as anxiety, depression, and poor quality of life (for review, see [5]). They also tend to be involved in risky behaviors such as smoking and to make poor lifestyle choices.

Interventions that enhance self-management skills with an emphasis on monitoring and managing health impacts can optimize physical and psychosocial functioning. Meta-analyses [6,7] show that face-to-face psychological interventions in pediatric chronic medical illnesses promote adherence and improve health outcomes. These interventions have large effect sizes irrespective of illness type, severity, and duration [8]. For treatment adherence, behavioral and multicomponent therapies have been shown to be more effective than educational programs alone [9]. The multicomponent approaches included combinations of a wide range of psychological therapies incorporating problem-solving, family behavioral therapy, biofeedback, social skills training, and psychosocial interventions.

Two recent studies [10,11] have investigated the efficacy of problem-solving skills training in improving well-being in children with a chronic respiratory condition (asthma) and from low-income families. For both, the training, delivered either at the individual family level [11] or as group lessons [10], involved face-to-face interactive problem-solving activities including defining problems, finding alternative solutions, and evaluating solutions. There was some support for the efficacy of the intervention with improved child academic performance and self-regulation [10] and health-related well-being [11] noted.

While face-to-face psychological interventions for adolescents with chronic health conditions are available in most hospitals, the limited availability and accessibility of such programs is a substantial barrier to more widespread uptake of routine psychosocial treatment or preventive interventions. The Internet offers a dynamic, interactive medium for providing information, changing attitudes and behavior, and enhancing social support [12]. It can create contexts where young people can feel safe enough to participate in activities and talk about issues important to them with other young people in similar situations. Despite their appeal and potential, Internet-based interventions for those with chronic health conditions are still quite novel. Stinson et al [13] systematically reviewed nine Internet interventions for pediatric conditions including asthma, pain, encopresis, obesity, and traumatic brain injury. They found positive results for Internet self-management interventions across a range of outcomes related to knowledge, behavioral change, and symptom management, but were unable to conclude that Internet interventions were effective in improving self-efficacy, social support, or emotional well-being. They also noted, however, that more rigorous randomized controlled trials were needed to endorse the use of Internet interventions.

In addition to the lack of robust randomized controlled trial data, many of these Internet-based interventions aimed to improve self-management only and did not provide peer social support. They did not address well-being in terms of increasing social support from others with a similar health condition or teaching skills for addressing life problems, but simply focused on problems associated with self-management. Increasing social support was a primary aim of one pilot study involving young people, aged between 13 and 18 years, with cystic fibrosis [14]. This Internet-based intervention consisted of an online support group including forums, email facility, and a graffiti wall. However, there were no chat facilities, so the communication between the participants was asynchronous. The adolescents in the study reported that they enjoyed emailing each other and going to the website, and were significantly more likely than before to respond that they had a friend with cystic fibrosis to whom they could relate following the intervention. Although promising, the findings of this study are limited, as it was not a randomized controlled trial and had no control group.

Present Study

Based on the finding that multicomponent interventions are more successful than educational interventions alone [9], our study incorporated social support and problem-solving skills to improve psychosocial well-being in children and adolescents with a chronic respiratory condition. The problem-solving paradigm was based on the work of D'Zurilla and Nezu [15], who framed problem solving as a cognitive–behavioral process whereby individuals endeavor to focus their coping efforts on altering the problematic nature of the situation, their reactions to the situation, or both. We developed an Internet intervention, Breathe Easier Online (BEO), specifically to improve social support and problem-solving skills. The intervention consisted of two parts: structured modules that the participants completed on a weekly basis to improve their social problem-solving skills; and an online community that incorporated both synchronous and asynchronous communication opportunities.

The aim of our study was to evaluate the efficacy of an online intervention for socially isolated children and adolescents with a chronic respiratory illness to improve psychosocial well-being. We hypothesized that, in this pilot randomized controlled trial, those receiving the BEO intervention would show improved

psychological well-being (as evidenced by lower depression scores) and problem-solving skills following intervention compared with those in a wait-list control group. We also anticipated positive secondary health-related outcomes for the BEO group, including an improved attitude toward their illness and improved treatment adherence.

Methods

Participants

We recruited 42 children aged between 10 and 17 years (mean age 13.58, SD 1.92 years) from the respiratory outpatient clinic at the Royal Children's Hospital, Brisbane, Australia. Children were eligible to participate if they spoke English as their primary language, had a primary diagnosis of a chronic respiratory condition, did not have any cognitive or sensory impairment that would preclude their completion of study measures, and were deemed socially isolated or disadvantaged by hospital staff based on social indicators (single parent who is unemployed or employed part-time, and known psychological and/or financial difficulties from hospital social work records). Exclusion criteria were children who were unable to use a computer, had an underlying psychiatric disorder, or had a recent (<3 months) hospitalization.

Ethical approval was granted by the Behavioral and Social Sciences Ethical Review Committee at the University of Queensland and the Human Ethics Committee of the Royal Children's Hospital, Brisbane. Informed consent was obtained from both the parent and participant in those aged 12 years or older and from parents alone in children aged under 12 years.

Protocol

Potential participants and their families were identified through hospital records by a research nurse at the hospital between July 2008 and November 2009, and were invited to participate. Following completion of the consent forms, children were randomly allocated to one of the two conditions: Internet-based intervention (BEO) or a wait-list control. Order of random allocation was predetermined via a computer program and was unknown to and concealed from the research staff. After baseline assessment (time 1) including completion of the package of questionnaires and spirometry, those in the BEO group received a Toshiba notebook computer and a modem to provide broadband Internet access for the duration of the study. Percentage of predicted forced expiratory volume in 1 second was measured with a spirometer that provides an objective measure of severity of obstructive lung diseases. The predicted value is based on Australian norms and depends on the child's height, age, and gender.

Paper-and-pencil questionnaire packages were completed at the hospital or at participants' homes at two time points (time 1: baseline/preintervention, and time 2: postintervention) for the BEO condition and at time equivalents to pre- and postintervention (9 weeks later) for the wait-list group. At time 2, the wait-list group was invited to participate in the intervention.

The Questionnaire Package

This consisted of several measures chosen to evaluate changes in the participants' psychological well-being (operationalized by a measure of depressive symptomatology), problem solving, attitudes toward their illness, and opinions and satisfaction with the intervention.

Center for Epidemiologic Studies Depression Scale for Children

The Center for Epidemiologic Studies Depression Scale for Children (CES-DC) [16] is a 20-item, self-report depression inventory with each item (eg, "During the past week, I felt like I was too tired to do things") rated on a 4-point Likert scale (0, not at all; 1, a little; 2, some; 3, a lot). Higher scores reflect higher depressive symptoms, with scores greater than 15 reflecting significant depressive symptoms [16]. The CES-DC has good internal consistency (alpha = .89) [16]. Cronbach alpha = .89 for the current sample.

Social Problem-Solving Inventory–Revised (Short Form)

The Social Problem-Solving Inventory–Revised (Short Form) (SPSI-R:SF) [17] comprises 25 items proposed to load onto 5 scales: positive problem orientation (eg, "Whenever I have a problem, I believe it can be solved"), negative problem orientation (eg, "I feel nervous and unsure of myself when I have an important decision to make"), rational problem solving (eg, "When I have a problem to solve, one of the first things I do is get as many facts about the problem as possible"), impulsive/careless style (eg, "I am too impulsive when it comes to making decisions"), and avoidance style (eg, "I go out of my way to avoid having to deal with problems in my life"). Items are rated on a 5-point Likert scale (0, not at all true of me; 1, slightly true of me; 2, moderately true of me; 3, very true of me; 4, extremely true of me) in which participants are asked to indicate how they would usually respond to problems. Higher scores on each scale indicate greater intensity on that attribute. The SPSI-R:SF has good internal consistency (alpha = .85 for adolescents) [17]. Cronbach alpha = .74 for the current sample.

Child Attitude Toward Illness Scale

The Child Attitude Toward Illness Scale (CATIS) [18] comprises 13 items designed to assess how favorably or unfavorably children feel toward having a chronic health condition (eg, "How often do you feel different from others because of your breathing problem?") on a 5-point Likert scale (1, never; 2, not often; 3, sometimes; 4, often; 5, very often). Higher scores reflect more positive attitudes toward illness. The CATIS has demonstrated good internal consistency (alpha = .77 to alpha = .89) and test–retest reliability over a 2-week interval (r = .77 to r = .80) in children and adolescents aged 8 to 17 years with epilepsy and asthma [18,19]. Cronbach alpha = .85 for the current sample.

Intervention Satisfaction Scale

This was a purpose-created self-report scale designed to explore the participants' opinions and satisfaction with the intervention, as well as seeking feedback at time 2. They rated 8 items (eg, "Would you recommend the program to others?") on a 4-point Likert scale (1, yes, very much so; 2, yes, for the most part; 3,

no, not really; 4, no, not at all). A final open-ended item allowed for any general comments or suggestions about the BEO intervention.

The BEO program

The interactive website for the participants in the BEO group (see Figure 1) was password protected and consisted of 5 components: My condition, My page, Daily diary, My work, and My talk.

My condition is a brief summary of each of the respiratory conditions that provided children with information about their own condition and conditions of the other children they encountered in the BEO program.

In *My page*, participants posted information about themselves including demographics, a photo (if they wished), favorite movie, favorite band, and a brief story about themselves. This page was visible to other BEO participants.

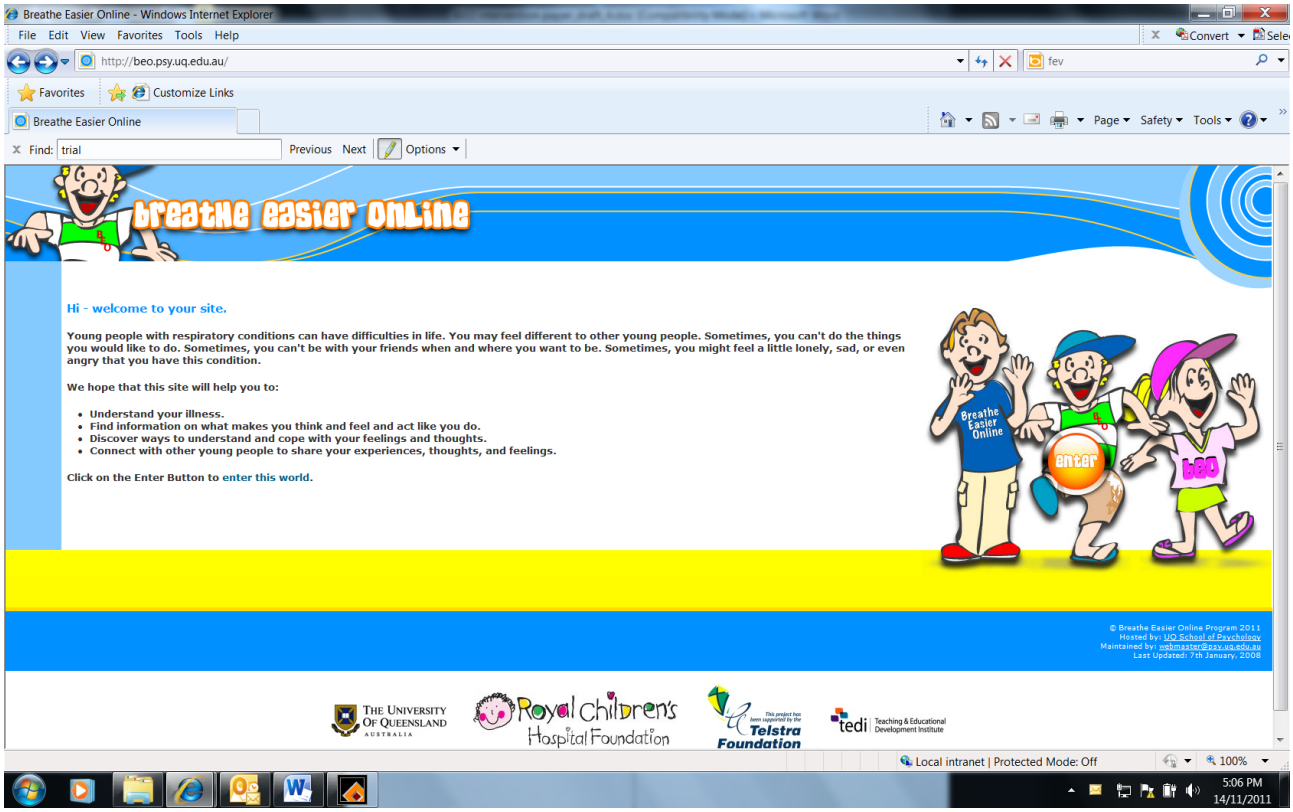
The *Daily diary* section contained a checklist where participants noted the medications they had taken each day. They also recorded how often they conversed with other participants in the program.

The *My work* section of the website contained the 6 modules that formed the focal intervention. The modules were based on D'Zurilla and Nezu's problem-solving theory [15] and provided interactive online skills training that targeted problems in general in addition to illness-specific problems. The online package followed a PACE principle for solving problems: problem identification, alternative solution generation, consequences of each alternative solution, execute solution and evaluate.

The *My talk* component of the website provided opportunities for BEO participants to communicate with each other. This communication could be either asynchronous (discussion board, email) or synchronous (instant messenger).

The 6 modules progressively and systematically introduced the participants to understanding their condition and treatment, the PACE program, and thinking processes. The intervention was interactive, as participants completed and submitted homework and were able to communicate online and regularly with research staff and other participants.

Figure 1. Screenshot of Breathe Easier Online home page.



Statistical Analysis

Sample size was based on an expected medium effect size of change in psychosocial measures for the intervention group with statistical power set at 90% and alpha =.05. We used SPSS version 17 (IBM Corporation, Somers, NY, USA) to analyze the data. An intention-to-treat approach to analyses was adopted with all participants included in the final analysis. The intention-to-treat analysis (missing values replaced with linear trend) ensured that those participants who completed time 1 measures and the intervention (but did not complete all time 2 measures) were included in the analyses. We used descriptive statistics to describe characteristics of the sample. Linear mixed models were used to assess the effects of the intervention on outcomes using a mixed-model analysis of variance (ANOVA) approach with time (time 1 vs time 2) as a within-subjects factor and group (BEO vs wait-list) as a between-subjects factor. Follow-up analyses were conducted in line with the stated hypotheses. The criterion for statistical significance was set at a 2-tailed value of $P \leq .05$.

Results

Characteristics of Participants

The flow diagram of those we approached and recruited, and who dropped out is depicted in Figure 2. Of the 42 participants who agreed to participate, 39 completed the trial, and 2 dropped out from the BEO intervention condition (final n = 19) and 1 from the wait-list control condition (final n = 20). All participants were white and 19/39 (49%) were male. Of the 39 participants, 12 (31%) had asthma, 22 (56%) had cystic fibrosis, 1 (3%) had tracheomalacia, and 2 (5%) had bronchiectasis.

Table 1 presents the data relating to demographic and respiratory conditions of the participants. There were no significant differences between the wait-list control and BEO intervention groups on any of the demographic, respiratory condition, or spirometry results at baseline (time 1; all $F < 1$). Of the 39 participants in the study, 35 were from single-parent families, with the majority (n = 20, 57%) not working. In addition, 13 families were living in rented accommodation and 10 families reported significant financial and/or psychological issues. A total of 12 of the children (31%) had reported poor adherence to treatment.

Table 1. Sample characteristics at baseline for both wait-list control and Breathe Easier Online (BEO) intervention conditions

	Wait-list control (n = 20)	BEO intervention (n = 19)	*P* value[a]
Age (years)			
Mean (SD)	13.63 (1.83)	13.41 (1.99)	.71
Gender			
Male/female	10/10	9/10	.87
Respiratory condition			
Cystic fibrosis	9	13	
Asthma	7	5	
Other	4	1	.24
Spirometry measures			
FEV_1[b], % predicted	74.19 (26.31)	78.17 (25.82)	.63
FVC[c], % predicted	85.10 (26.52)	89.37 (26.76)	.61
$FEF_{25\%–75\%}$ (L/min)[d]	61.32 (27.72)	62.71 (27.28)	.87

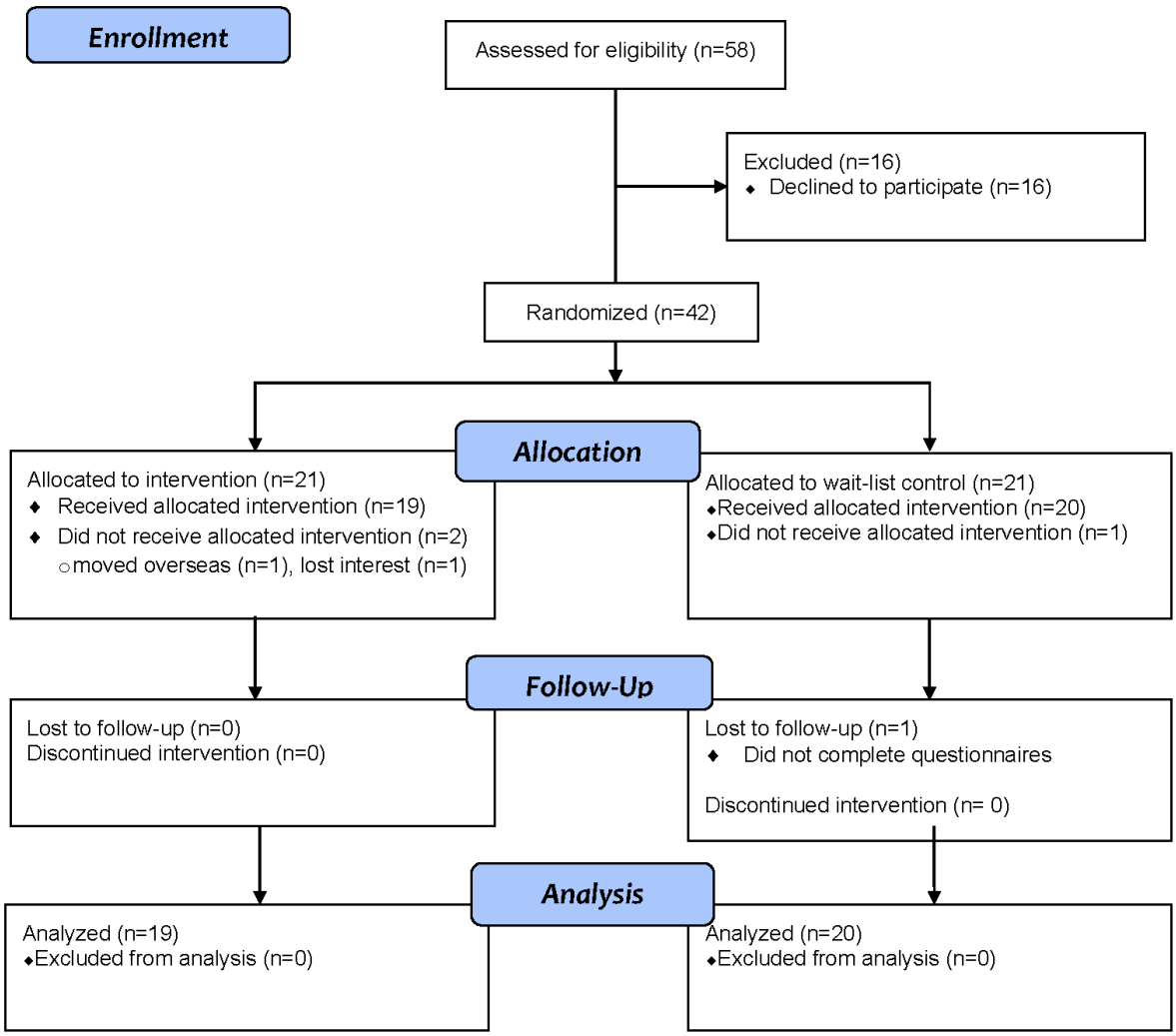
[a] *P* values for *t* tests of group differences.

[b] Forced expiratory volume in 1 second.

[c] Forced vital capacity.

[d] Forced expiratory flow, midexpiratory phase.

Figure 2. CONSORT flow diagram for the randomized controlled trial of the Breathe Easier Online intervention.



Effect of BEO Intervention

The descriptive data for the mixed-models ANOVAs are presented in Table 2, showing primary outcomes of depressive symptoms and social problem-solving skills, and secondary outcomes of attitude toward illness and spirometry (as an indicator of treatment adherence). The results for the analyses of depression scores revealed neither significant group ($F_{1,33}$ = 2.00, P = .17, η^2 = .06) or time (F < 1, P = .35, η^2 = .03) main effects, nor a significant interaction ($F_{1,33}$ = 1.76, P = .19, η^2 = .05). Similar results were found for full scale scores on social problem solving with neither significant main effects for group (F < 1, P = .61, η^2 = .01) or time ($F_{1,31}$ = 2.00, P = .17, η^2 = .06), nor their interaction ($F_{1,31}$ = 1.82, P = .19, η^2 = .06). These nonsignificant main and interaction effects were repeated for each of the social problem subscales except for impulsive/carelessness, where a significant time main effect was evident ($F_{1,31}$ = 9.72, P = .004, η^2 = .24). Irrespective of participation in the intervention, the young people reported less of this problem-solving behavior at time 2 (mean 7.06, SD 3.43) than they did at time 1 (mean 9.18, SD 4.58). The mixed-model ANOVA on participants' attitudes toward illness resulted in nonsignificant main effects of time ($F_{1,33}$ = 3.23, P = .08, η^2 = .09) and group (F < 1, η^2 = .01), as well as their interaction ($F_{1,33}$ = 1.26, P = .27, η^2 = .04).

For the spirometry results, both main effects of group ($F_{1,35}$ = 1.69, P = .20, η^2 = .05) and time (F < 1, P = .65, η^2 = .01) were nonsignificant, as was their interaction (F < 1, P = .64, η^2 = .01).

Despite these nonsignificant findings, we conducted further analyses to explore more fully and specifically the hypotheses of interest. These analyses incorporated examination of group differences on participants' change scores (ie, difference in scores between time 1 and time 2), differences across time for each of the BEO and wait-list groups separately, and a priori levels of clinically significant (as opposed to statistically significant) change.

The findings from the analyses of change scores are summarized in Table 3. There were no significant group differences for change scores across all outcome measures. However, further analyses investigating the significance of the differences across time for the BEO and wait-list groups separately did identify some noteworthy findings. In line with the stated hypotheses, the depression scores for the BEO group following intervention (mean 9.53, SD 7.54) were significantly lower than reported at baseline (mean 13.67, SD 10.52; $F_{1,33}$ = 4.38, P = .04, η^2 = .12). Such was not the case for the wait-list group (time 1: mean 16.41, SD 15.64; time 2: mean 17.09, SD 13.24; F < 1, P = .54, η^2 = .01).

For the SPSI-R:SF, follow-up analyses indicated that those in the BEO group were significantly less likely to report an impulsive/careless problem-solving style following their intervention (mean 5.88, SD 2.89) than at time 1 (mean 8.44, SD 4.10; $F_{1,31}$ = 6.80, P = .01, η^2 = .18). The wait-list group showed no significant difference in this approach to problem solving across the same time period (time 1: mean 9.88, SD 5.01; time 2: mean 8.18, SD 3.61; $F_{1,31}$ = 3.20, P = .08, η^2 = .09).

Further, the participants in the BEO group reported a significantly better attitude toward their illness following the intervention (mean 3.61, SD 0.60) than at baseline (mean 3.29, SD 0.64; $F_{1,33}$ = 4.40, P = .04, η^2 = .12). There was no significant difference in the wait-list group (F < 1, P = .64, η^2 = .01).

In some settings, differences between groups may be considered of clinical significance despite nonsignificant statistical differences. A common decision criterion for clinical significance is a change in scores (or difference between the groups) equal to one-half of the average of the pooled standard deviation [20]. Based on this criterion, there was evidence of trends toward clinical significance in group differences for depression, impulsive/careless problem-solving style, and attitudes toward illness with improvements supporting the BEO group.

We analyzed responses to the Intervention Satisfaction Scale questionnaire to gauge the participants' perceptions of the BEO program following their completion of the online modules. Almost all (18/19, 95% of participants) reported that they were happy to do the program and "thoroughly enjoyed" it (15/19, 79%). Most (18/19, 95%) stated that they would highly recommend the program to others. Only 2 participants dropped out of the program (1 family relocated) and all participants who remained in the project completed all 6 online modules.

Table 2. Descriptive statistics on outcome measures for children and adolescents for Breathe Easier Online (BEO) intervention and wait-list control conditions

Outcome measure	BEO Intervention		Wait-list	
	Mean	SD	Mean	SD
CES-DC [a]				
Time 1	13.67	10.52	16.41	15.64
Time 2	9.53	7.54	17.09	13.24
SPSI-R:SF [b]				
Positive problem orientation				
Time 1	10.44	4.44	11.35	3.14
Time 2	10.28	3.74	10.65	4.66
Negative problem orientation				
Time 1	5.88	4.90	5.82	2.88
Time 2	4.06	3.09	6.85	5.68
Rational problem solving				
Time 1	8.06	4.42	9.24	4.91
Time 2	8.82	5.93	8.88	3.68
Impulsive/careless style				
Time 1	8.44	4.10	9.88	5.01
Time 2	5.88	2.89	8.18	3.61
Avoidance style				
Time 1	7.19	4.71	7.06	3.17
Time 2	6.13	4.43	6.48	5.03
Overall score				
Time 1	11.40	3.29	11.56	1.86
Time 2	12.62	2.65	11.59	2.86
CATIS [c]				
Time 1	3.29	0.64	3.31	0.76
Time 2	3.61	0.60	3.38	0.72
FEV1pred[d] (%)				
BEO	83.17	18.88	74.19	26.31
Wait-list	83.19	17.20	71.98	32.52

[a] Center for Epidemiologic Studies Depression Scale for Children (higher scores reflect greater depressive symptomatology).

[b] Social Problem-Solving Inventory–Revised (Short Form) (higher scores reflect greater use of that problem-solving style).

[c] Child Attitude Toward Illness Scale (higher scores reflect a more positive attitude toward the illness).

[d] Predicted forced expiratory volume in 1 second.

Table 3. Descriptive statistics on change scores (time 1 to time 2 differences) of outcome measures for children and adolescents for Breathe Easier Online (BEO) intervention and wait-list control conditions

Outcome measure	Change scores				*P* value[a]	Effect size
	BEO Intervention		Wait-list			
	Mean	SD	Mean	SD		
CES-DC[b]	–4.14	7.56	0.68	13.29	.19	.05
SPSI-R:SF [c]						
Positive problem orientation	–0.16	4.00	–0.71	5.38	.74	.01
Negative problem orientation	–1.81	4.20	1.03	4.94	.09	.09
Rational problem solving	0.81	4.15	–0.41	3.32	.36	.03
Impulsive/careless style	–2.56	3.16	–1.71	4.54	.54	.01
Avoidance style	–1.06	5.08	–0.59	5.40	.79	.01
Overall score	1.22	2.62	0.03	2.45	.19	.06
CATIS[d]	0.32	0.57	0.07	0.73	.27	.04
FEV_1pred[e]	0.03	13.47	–2.21	15.19	.64	.01

[a] *P* value comparing group differences on change scores.

[b] Center for Epidemiologic Studies Depression Scale for Children (negative scores represent fewer symptoms of depression at time 2 than at time 1).

[c] Social Problem-Solving Inventory (negative scores represent less of that style at time 2 than at time 1).

[d] Child Attitude Toward Illness Scale (positive change scores represent a more positive attitude at time 2 than at time 1).

[e] Predicted forced expiratory volume in 1 second.

Discussion

BEO is an Internet-based intervention designed to improve the well-being of children and adolescents with a chronic respiratory condition. Although there were no significant group differences following the intervention, there was evidence, albeit preliminary and tentative, of the efficacy of the program in improving attitudes toward illness, reducing depression symptoms, and decreasing maladaptive social problem solving (impulsive/careless style) for the participants in the BEO group following their intervention. The program was met with considerable enthusiasm from the participants and their parents. The majority of families reported they would definitely recommend it to others. The attrition rate was very low, as participants enjoyed interacting on the specifically created website.

Comparisons with Prior Work

The present study extends on previous research [21-24] by demonstrating the initial efficacy and feasibility of Internet-based interventions with young people with chronic health conditions. However, many of the previously reported interventions have not specifically addressed the young person's psychosocial well-being but have focused on symptom monitoring [21], pain treatment [25], health care [26], and education [23]. In fact, where psychosocial well-being was an outcome, the findings have been equivocal [21,22]. Although acknowledging the methodological, intervention, and illness differences that exist between the present study and this past research, it is important to note that our study is unique in that we found modest improvements in problem solving and attitudes toward illness and lowering depression following our online intervention.

Limitations

Despite the promising findings, our study has some limitations that must be recognized. The nonsignificant group differences across all outcome measures need to be interpreted in light of several factors. It may be that the intervention, itself, was not sufficiently focused or presented on a platform that would engage the participants in a way that would lead to hypothesized improvements. While necessarily directed at set and established problems, the online modules might have provided more room for participant-own identification of their problems followed by therapist-guided group identification of resolutions. This would have created a greater relevance, engagement, and interaction with the modules. The sample size was small, and this may have affected our ability to detect true group differences across other outcome measures (ie, statistical power). Although statistically nonsignificant, some of our findings were encouraging, indicating trends that a larger sample size might lead to significant results. Another limitation concerned the participants' baseline well-being scores. These were not at clinical levels (except for the depression scores in the wait-list group), and this may have hindered the likelihood of improvements showing following the intervention. There is some evidence that interventions for participants not within a clinical range do not produce changes of significance in symptoms [27]. Finally, the children progressed unevenly and at various speeds through the online modules with life events (eg, holidays) disrupting progress. This is an acknowledged problem with Internet interventions with children [28] and can have an influence on their impact.

Future Research

This study has provided pilot data for future work, as the small sample size renders this study underpowered. Also, the findings of the current pilot study highlight several avenues for future research. Despite the participants' engagement with the Web-based modules, anecdotal feedback suggests that they were too text intensive. Future research may investigate the efficacy of Internet-based interventions using different modalities (text, picture, or video) across different age groups to find the optimal age-related combination. While the children and adolescents in our BEO group demonstrated positive gains, the question as to whether these gains compare favorably with those of face-to-face interventions remains unanswered, as does the question of the maintenance of any gains in well-being. Future research that makes direct comparisons with a face-to-face intervention group is necessary to ensure that the online environment does overcome some of the face-to-face intervention barriers (eg, accessibility) but not at the expense of gains in well-being.

Conclusions

The Internet offers an exciting possibility for intervention with children and adolescents with chronic respiratory conditions, as it can be interactive, engaging, and fun and can provide unique opportunities for online peer and professional support that can continue once the intervention has been completed. Moreover, it can overcome some of the barriers and impediments found in face-to-face interventions. Our findings are promising but also highlight the need for further research attention toward specifically designed online programs as an intervention with children and adolescents with chronic health conditions.

Acknowledgments

This work was supported by the Telstra Foundation and Royal Children's Hospital Foundation (Grant Number 10237). AC is funded by an NHMRC Practitioner Fellowship (Grant Number 545216).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (V1.6) checklist [29].

[PDF File (Adobe PDF File), 814KB - jmir_v14i1e23_app1.pdf]

References

1. Australian Bureau of Statistics. National Health Survey: Summary of Results, 2007-2008. Canberra: ABS; 2009. URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4364.0Main+Features12007-2008%20(Reissue)?OpenDocument [accessed 2012-01-30] [WebCite Cache ID 655muK3HC]
2. Almqvist C, Pershagen G, Wickman M. Low socioeconomic status as a risk factor for asthma, rhinitis and sensitization at 4 years in a birth cohort. Clin Exp Allergy 2005 May;35(5):612-618. [doi: 10.1111/j.1365-2222.2005.02243.x] [Medline: 15898983]
3. Cesaroni G, Farchi S, Davoli M, Forastiere F, Perucci CA. Individual and area-based indicators of socioeconomic status and childhood asthma. Eur Respir J 2003 Oct;22(4):619-624 [FREE Full text] [Medline: 14582914]
4. Bush JP, Huchital JR, Simonian SJ. An introduction to program and research initiatives of the Starbright Foundation. Child Health Care 2002;31(1):1-10. [doi: 10.1207/S15326888CHC3101_1]
5. McQuaid EL, Kopel SJ, Nassau JH. Behavioral adjustment in children with asthma: a meta-analysis. J Dev Behav Pediatr 2001 Dec;22(6):430-439. [Medline: 11773808]
6. Beale IL. Scholarly literature review: Efficacy of psychological interventions for pediatric chronic illnesses. J Pediatr Psychol 2006 Jun;31(5):437-451. [doi: 10.1093/jpepsy/jsj079] [Medline: 16162841]
7. Kibby MY, Tyc VL, Mulhern RK. Effectiveness of psychological intervention for children and adolescents with chronic medical illness: a meta-analysis. Clin Psychol Rev 1998 Jan;18(1):103-117. [Medline: 9455625]
8. Northam EA, Matthews LK, Anderson PJ, Cameron FJ, Werther GA. Psychiatric morbidity and health outcome in Type 1 diabetes--perspectives from a prospective longitudinal study. Diabet Med 2005 Feb;22(2):152-157. [doi: 10.1111/j.1464-5491.2004.01370.x] [Medline: 15660731]
9. Kahana S, Drotar D, Frazier T. Meta-analysis of psychological interventions to promote adherence to treatment in pediatric chronic health conditions. J Pediatr Psychol 2008 Jul;33(6):590-611. [doi: 10.1093/jpepsy/jsm128] [Medline: 18192300]
10. Clark NM, Shah S, Dodge JA, Thomas LJ, Andridge RR, Little RJ. An evaluation of asthma interventions for preteen students. J Sch Health 2010 Feb;80(2):80-87. [doi: 10.1111/j.1746-1561.2009.00469.x] [Medline: 20236406]
11. Seid M, Varni JW, Gidwani P, Gelhard LR, Slymen DJ. Problem-solving skills training for vulnerable families of children with persistent asthma: report of a randomized trial on health-related quality of life outcomes. J Pediatr Psychol 2010 Nov;35(10):1133-1143. [doi: 10.1093/jpepsy/jsp133] [Medline: 20061311]

12. Bull SS, Gaglio B, McKay HG, Glasgow RE. Harnessing the potential of the internet to promote chronic illness self-management: diabetes as an example of how well we are doing. Chronic Illn 2005 Jun;1(2):143-155. [Medline: 17136920]
13. Stinson J, Wilson R, Gill N, Yamada J, Holt J. A systematic review of internet-based self-management interventions for youth with health conditions. J Pediatr Psychol 2009 Jun;34(5):495-510. [doi: 10.1093/jpepsy/jsn115] [Medline: 19029142]
14. Johnson KB, Ravert RD, Everton A. Hopkins Teen Central: Assessment of an internet-based support system for children with cystic fibrosis. Pediatrics 2001 Feb;107(2):E24. [Medline: 11158498]
15. D'Zurilla TJ, Nezu AM. Problem-Solving Therapy: A Social Competence Approach to Clinical Intervention. 2nd edition. New York, NY: Springer; 1999.
16. Fendrich M, Weissman MM, Warner V. Screening for depressive disorder in children and adolescents: validating the Center for Epidemiologic Studies Depression Scale for Children. Am J Epidemiol 1990 Mar;131(3):538-551. [Medline: 2301363]
17. D'Zurilla TJ, Nezu AM, Maydeu-Olivares A. SPSI-R: Social Problem-Solving Inventory-Revised (SPSI-R). North Tonawanda, NY: Multi-Health Systems; 2002.
18. Austin JK, Huberty TJ. Development of the Child Attitude Toward Illness Scale. J Pediatr Psychol 1993 Aug;18(4):467-480. [Medline: 8410570]
19. Heimlich TE, Westbrook LE, Austin JK, Cramer JA, Devinsky O. Brief report: Adolescents' attitudes toward epilepsy: further validation of the Child Attitude Toward Illness Scale (CATIS). J Pediatr Psychol 2000;25(5):339-345. [Medline: 10880064]
20. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: L Erlbaum Associates; 1988.
21. Chan DS, Callahan CW, Hatch-Pigott VB, Lawless A, Proffitt HL, Manning NE, et al. Internet-based home monitoring and education of children with asthma is comparable to ideal office-based care: results of a 1-year asthma in-home monitoring trial. Pediatrics 2007 Mar;119(3):569-578 [FREE Full text] [doi: 10.1542/peds.2006-1884] [Medline: 17332210]
22. Jan RL, Wang JY, Huang MC, Tseng SM, Su HJ, Liu LF. An internet-based interactive telemonitoring system for improving childhood asthma outcomes in Taiwan. Telemed J E Health 2007 Jun;13(3):257-268. [doi: 10.1089/tmj.2006.0053] [Medline: 17603828]
23. Runge C, Lecheler J, Horn M, Tews JT, Schaefer M. Outcomes of a Web-based patient education program for asthmatic children and adolescents. Chest 2006 Mar;129(3):581-593 [FREE Full text] [doi: 10.1378/chest.129.3.581] [Medline: 16537855]
24. Stinson JN, McGrath PJ, Hodnett ED, Feldman BM, Duffy CM, Huber AM, et al. An internet-based self-management program with telephone support for adolescents with arthritis: a pilot randomized controlled trial. J Rheumatol 2010 Sep;37(9):1944-1952. [doi: 10.3899/jrheum.091327] [Medline: 20595280]
25. Hicks CL, von Baeyer CL, McGrath PJ. Online psychological treatment for pediatric recurrent pain: a randomized evaluation. J Pediatr Psychol 2006 Aug;31(7):724-736. [doi: 10.1093/jpepsy/jsj065] [Medline: 16093516]
26. Joseph CL, Peterson E, Havstad S, Johnson CC, Hoerauf S, Stringer S, Asthma in Adolescents Research Team. A web-based, tailored asthma management program for urban African-American high school students. Am J Respir Crit Care Med 2007 May 1;175(9):888-895 [FREE Full text] [doi: 10.1164/rccm.200608-1244OC] [Medline: 17290041]
27. Sawyer MG, Pfeiffer S, Spence SH, Bond L, Graetz B, Kay D, et al. School-based prevention of depression: a randomised controlled study of the beyondblue schools research initiative. J Child Psychol Psychiatry 2010 Feb;51(2):199-209. [doi: 10.1111/j.1469-7610.2009.02136.x] [Medline: 19702662]
28. 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] [Medline: 18794187]
29. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ANOVA: analysis of variance
BEO: Breathe Easier Online
CATIS: Child Attitude Toward Illness Scale
CES-DC: Center for Epidemiologic Studies Depression Scale for Children
FEV1: forced expiratory volume in 1 second
PACE: problem, alternative solutions, consequences, execute
SPSI-R:SF: Social Problem-Solving Inventory–Revised (Short Form)

Edited by G Eysenbach; submitted 16.11.11; peer-reviewed by T Palermo, N Zwar; comments to author 06.12.11; revised version received 18.12.11; accepted 13.01.12; published 08.02.12.

Please cite as:
Newcombe PA, Dunn TL, Casey LM, Sheffield JK, Petsky H, Anderson-James S, Chang AB
Breathe Easier Online: Evaluation of a Randomized Controlled Pilot Trial of an Internet-Based Intervention to Improve Well-being in Children and Adolescents With a Chronic Respiratory Condition
J Med Internet Res 2012;14(1):e23
URL: http://www.jmir.org/2012/1/e23/
doi:10.2196/jmir.1997
PMID:22356732

Original Paper

Effects of a Web-Based Intervention for Adults With Chronic Conditions on Patient Activation: Online Randomized Controlled Trial

Michael Solomon[1], MBA, PhD; Stephen L Wagner[2], MS, PhD; James Goes[3], PhD

[1]Point-of-Care Partners, Coral Springs, FL, United States

[2]Division of Medical Education and Research, Carolinas Healthcare System, Charlotte, NC, United States

[3]Department of Health Sciences, Walden University, Minneapolis, MN, United States

Corresponding Author:
Michael Solomon, MBA, PhD
Point-of-Care Partners
11236 NW 49 Street
Coral Springs, FL, 33076
United States
Phone: 1 561 400 7091
Fax: 1 561 400 7091
Email: michael.solomon@pocp.com

Abstract

Background: With almost one-half of Americans projected to have at least one chronic condition before 2020, a vital role of the health care system is to develop informed, engaged individuals who are effective self-managers of their health. Self-management interventions (SMIs) delivered face-to-face or by telephone (traditional SMIs) are associated with improved self-management knowledge, skills, and self-efficacy, which are expressed by the composite construct of patient activation, a predictor of health outcomes. Web-based interventions to support self-management across the spectrum of chronic diseases have the potential to reach a broader population of patients for extended periods than do traditional SMIs. However, evidence of the effectiveness of Web-based interventions on patient activation is sparse. High-quality studies featuring controlled comparisons of patients with different chronic conditions are needed to explore the interaction of Web-based interventions and patient activation.

Objective: To explore the effect of a Web-based intervention on the patient activation levels of patients with chronic health conditions, measured as attitudes toward knowledge, skills, and confidence in self-managing health.

Methods: For this 12-week study, prospective participants were selected from the patient panel of a regional health care system in the United States. The 201 eligible participants were randomly assigned to two groups. Intervention group participants had access to MyHealth Online, a patient portal featuring interactive health applications accessible via the Internet. Control participants had access to a health education website featuring various topics. Patient activation was assessed pre- and posttest using the 13-item patient activation measure. Parametric statistical models (*t* test, analysis of variance, analysis of covariance) were applied to draw inferences.

Results: The Web-based intervention demonstrated a positive and significant effect on the patient activation levels of participants in the intervention group. A significant difference in posttest patient activation scores was found between the two groups ($F_{1,123}$ = 4.438, P = .04, r = .196). Patients starting at the most advanced development of patient activation (stage 4) in the intervention group did not demonstrate significant change compared with participants beginning at earlier stages.

Conclusions: To our knowledge, this is the first study to measure change in patient activation when a Web-based intervention is used by patients living with different chronic conditions. Results suggest that Web-based interventions increase patient activation and have the potential to enhance the self-management capabilities of the growing population of chronically ill people. Activated patients are more likely to adhere to recommended health care practices, which in turn leads to improved health outcomes. Designing Web-based interventions to target a specific stage of patient activation may optimize their effectiveness. For Web-based interventions to reach their potential as a key component of chronic disease management, evidence is needed that this technology produces benefits for a sustained period among a diverse population.

(J Med Internet Res 2012;14(1):e32) doi:10.2196/jmir.1924

KEYWORDS

Chronic care; health information technology; patient activation; randomized controlled experiment; self-management; Web-based intervention

Introduction

Care for patients with chronic diseases consumes 78% of the total cost of a US health care system that has been slow to adapt to the complex needs of this growing population [1,2]. With almost one-half of Americans projected to have at least one chronic condition before the end of this decade [3,4], a vital role of the health care system is to provide the tools necessary for chronically ill patients to make informed decisions about their health care, and to solve the problems encountered daily from living with a chronic condition [5,6]. Self-management programs are designed to aid in this development by educating patients about their diseases, teaching skills to promote self-care behaviors, and fostering self-confidence in patients' abilities to manage their disease [7-9]. As patients' capabilities in these three areas improve, their level of patient activation, a measure of self-management capabilities, increases [10]. This paper describes the results of a study of change in patient activation when a self-management program delivered via the Internet is used. The findings inform innovators of chronic care programs on strategies for leveraging information technology to address the challenges of delivering self-management support services to a large and growing population of patients.

Patients who believe they have a responsibility to take an active role in making decisions about their health are central to effective chronic care management [11,12]. As the principal managers of their own care [5], activated patients are more likely to adhere to activities for controlling symptoms and the progression of their disease [13,14]. A person's activation level indicates the extent of his or her self-management capabilities, encompassing knowledge, skills, and self-efficacy [10,15]. Development of these self-management capabilities precedes change in health behavior or status [16,17]. Thus, the construct of patient activation is a predictor of health process and outcomes measures [14,18] and is therefore a key indicator of the quality of chronic care management [10].

Activated patients strive to understand their conditions, viewing problems as challenges with the confidence that they can be solved [8,19]. Acquiring these attributes of self-management is a learning process, involving the acquisition of knowledge and problem-solving skills that enable an individual to confidently engage in decision making and actions to effectively manage their chronic health condition [13]. The developmental nature of patient activation is represented by four stages. At the earliest stage, individuals form beliefs that taking control of their health is important. As people progress through the second and third stages of patient activation, they develop the knowledge and skills to become increasingly active in self-managing their health. The most activated patients (stage 4) sustain self-management behaviors except when confronted with new or stressful situations [19]. Considering that patients are confronting different types of challenges at each stage of patient activation, researchers hypothesize that the most effective self-management interventions (SMIs) will feature designs targeting patients at a particular stage [13]. Despite the implications for designers of Web-based interventions for self-management, no research testing this theory appears in the literature.

Evidence suggests that SMIs delivered by face-to-face and telephonic modalities (ie, traditional SMIs) are associated with improvements in self-management knowledge and skills [17,20] and self-efficacy [21] among patients with various chronic conditions. Traditional SMI studies measuring change in the composite of these self-management capabilities—expressed as patient activation—show improvement in patient activation levels of chronically ill participants [13]. Thus, traditional SMIs demonstrate effectiveness in helping patients develop their self-management capabilities and therefore serve a vital role in the broader goal of improving health outcomes.

Use of information technology to deliver SMIs via the Internet has the potential to reach a broader population of chronically ill patients for extended periods of time when compared with traditional SMIs [9,22,23]. Web-based interventions (also referred to as Internet-based interventions) are applications accessed via a website by patients and are designed to improve understanding of a health condition, change health behavior, and enhance problem-solving skills [24,25]. Certain Web-based interventions are aimed at helping patients develop self-management capabilities and modify self-care behaviors to better deal with their chronic conditions [23,26,27]. The ensuing discussion of Web-based interventions refers to applications with self-management support as the central focus.

The health care community's understanding of the value of Web-based interventions is inhibited by a dearth of high-quality studies [28], high variability in the effectiveness of different types of Web-based interventions [24], and mixed results of their effect on self-efficacy [28,29], a key component of patient activation. Few studies of Web-based interventions have explored the broader construct of patient activation (encompassing knowledge, skills, and self-efficacy), an area of inquiry that holds promise because of the symbiotic relationship between knowledge, skills, and self-efficacy in self-management performance [8,16]. Furthermore, the literature on Web-based intervention research is dominated by disease-specific interventions and measures, with scant evidence of the effectiveness of applications designed for use by populations with a broad spectrum of chronic diseases. More experimental studies featuring controlled comparisons of samples comprising patients with different chronic conditions are needed to better understand the interaction between Web-based interventions and the attributes of self-management performance constituting patient activation [30,31].

The few high-quality studies in the literature of Web-based interventions and their effectiveness in developing self-management capabilities suggest that these applications help patients to understand their role in managing health and the fundamental aspects of their chronic conditions. For

example, Web-based interventions increase patients' awareness of the need to be actively engaged in their health care [32]. A systematic review of randomized controlled trials revealed that Web-based interventions with self-management education modules had a significant and positive effect on patients' knowledge of their chronic conditions [29], a component of patient activation.

Research on the effects of Web-based interventions on the self-efficacy of people with chronic conditions is limited and results are mixed. Although the effect of Web-based interventions on self-efficacy showed promise in a meta-analysis [29], an insufficient number of studies precluded any conclusions. Web-based interventions designed specifically for patients with chronic conditions such as diabetes or arthritis showed significant enhancement in self-efficacy levels [28,33,34] compared with a more generalized application targeting broader chronic disease populations, which did not demonstrate a significant effect [35]. Although more research is needed before conclusions should be drawn, these results suggest that Web-based interventions designed to target specific chronic diseases may increase self-efficacy; no evidence exists that applications targeting a diverse population of chronically ill patients influence this important element of patient activation.

Only one study of a Web-based intervention's effect on the composite construct of patient activation of patients with a chronic condition appears in the literature. This randomized controlled trial reported a significant improvement in patient activation among patients with diabetes [33]. The study described in this paper builds on that research by evaluating the effect of a Web-based intervention on patient activation of patients living with different chronic conditions, including asthma, diabetes, and hypertension.

The aim of this randomized controlled trial was to evaluate change in self-management capabilities—expressed as patient activation—when a group of chronically ill patients were provided with online access to self-management materials. Specifically, we hypothesized that patients given access to a Web-based intervention designed to support self-management in the context of a person's particular chronic disease would demonstrate positive change in patient activation levels compared with control group participants. A secondary hypothesis was tested within the intervention group. We hypothesized that patients beginning at a lower stage of patient activation development would demonstrate greater change in patient activation than would participants starting at higher stages.

Methods

Study Setting and Participants

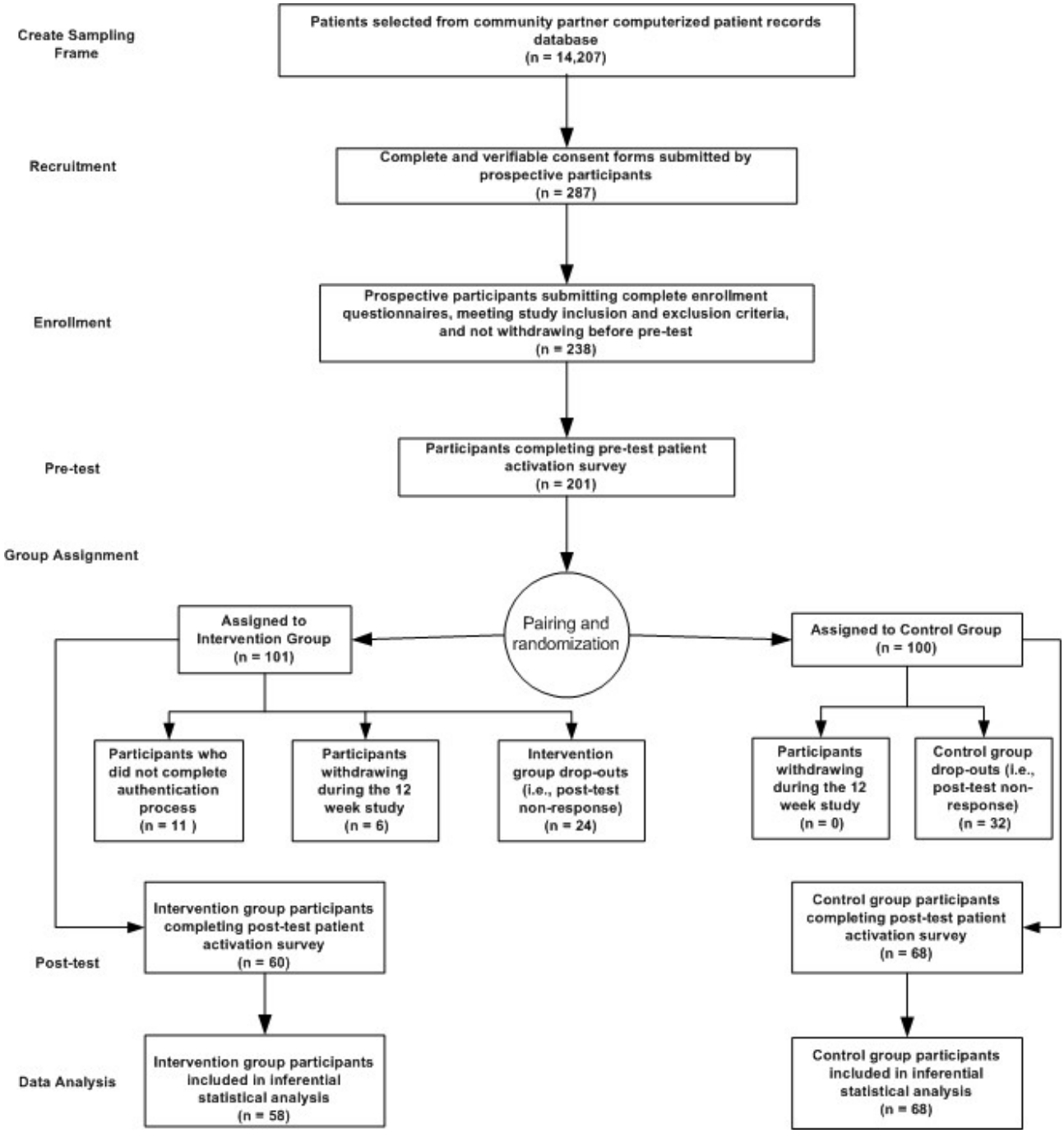
Prospective participants were selected from the patient panel of Carolinas HealthCare System, a regional health care-delivery system in the southeastern United States. Adult patients who were seen by 300 physicians employed by Carolinas HealthCare System and with a diagnosis of asthma, hypertension, or diabetes constituted the sampling frame. We selected patients between ages 18 and 64 years, inclusive, with a diagnosis of one of the three conditions, and who had visited a participating physician in the past 2 years but not in at least 180 days. Patients with a chronic disease who have not visited their doctor in at least 6 months suggests nonadherence with recommended care guidelines and may indicate low patient activation. Employees of Carolinas HealthCare System were excluded. Selected patients were sent personalized invitations to participate. This study's research protocol was approved by the Carolinas HealthCare System Research Review Committee, which did not require an external trial registration for this health behavior research. The institutional review boards of Walden University and Carolinas HealthCare System also approved the study.

All information used in this study for group assignments and data analysis was self-reported by participants and collected using a Web-based survey tool. Figure 1 shows the participants' flow through the 12-week study.

Interested patients submitted online consents and enrollments. Patients meeting the eligibility criteria were placed in the pool of participants for the study. Participants were randomly assigned to the intervention or control group from pairs created by matching on adherence scores. We used a matching process because it helps to mitigate differences between the groups at baseline and strengthens the study's statistical power [36]. The adherence scores were calculated based on participants' responses to four items in the enrollment questionnaire related to a person's adherence to self-management behavior. The two participants with the lowest adherence scores constituted the first pair and those with the highest scores, the last. Starting at an arbitrary point in the stack of pairs, we assigned one member of the pair to the intervention and the other to the control group.

On completion of the random assignments, participants were notified via express mail of their inclusion in the study, were presented with descriptive information regarding the self-management material that would be made available to them, and received instructions for accessing the pretest and study material on the Internet. Included in these directions was a unique and confidential login that controlled the specific participant's access privileges to the appropriate intervention or control group material. Participants were not informed of their intervention or control status.

Figure 1. Flow of study participants.



Materials

We measured patient activation using the 13-item patient activation measure (PAM-13) [37]. The PAM-13 is designed to elicit responses from a person about his or her attitudes toward knowledge, skills, and confidence in self-managing health. The scale is based on the Guttman technique with items ordered according to level of difficulty. A 4-point Likert-type scale of response options ranging from strongly disagree to strongly agree is used to elicit endorsement of a particular statement. The PAM-13 statements are published by the developers elsewhere [37]. PAM-13 item responses result in total raw scores ranging from 13 to 52, which we converted to the linear interval scale of patient activation scores, ranging from 0 (lowest activation) to 100 (highest activation) [38]. Within this converted scale are cut-off minimum-point levels for each of the four stages of patient activation described earlier [13,19].

Psychometric analysis of the PAM-13 reported in the literature shows a measure with strong reliability and validity properties [37]. The unaltered PAM-13 was the pretest survey. The posttest survey included the PAM-13 plus an additional question about the participant's visits to doctors during the study period.

The Web-based intervention used in this study was MyHealth Online, a personal health portal featuring a suite of interactive health applications. Effective Web-based interventions feature functions designed to change health behavior and improve patient–provider communications [24]. The MyHealth Online self-service and health education applications enable patients

to manage their health care directly. MyHealth Online users can book doctors' office appointments online, request prescription renewals, and view and pay their bills.

MyHealth Online's interactive, multimedia health education modules are based on information therapy principles, with each online session designed to advance the user's knowledge by providing evidence-based information on the patient's specific condition, self-management guidelines, and options for problem solving and treatment [39]. Each week of the study, intervention participants received an email alerting them to the availability of the next in a progressive series of health education sessions specific to their chronic condition. Patients interact with the MyHealth Online health education modules at their own pace and decide the level of complexity of the material for a particular session. The self-directed design and 24-hour per day availability of MyHealth Online ensures that the information a patient needs to manage a particular problem is available when it is needed, facilitating decision making and behavior change [39].

If patients need more information, they can communicate online with their providers using the secure message function of MyHealth Online. Secure provider–patient email communication helps patients engage in their care and improves their access to information [40]. The software supporting MyHealth Online is supplied by GE Healthcare (Waukesha, WI, USA) and the health education applications are provided by Healthwise (Boise, ID, USA).

Control group participants were provided with access to a website hosting health education material on a variety of topics. In contrast to MyHealth Online, the materials available to control group participants were noninteractive and not prescriptive like the health education material provided to the intervention group. Participants were required to search topics to locate content of interest. Access to MyHealth Online or the control group's website was granted to participants in the respective groups after they completed the pretest survey. All participants in the control and intervention groups were encouraged to continue their usual care during the study.

All participants in the study had access to the Carolinas HealthCare System website help desk for guidance on accessing and using the applications and to resolve technical problems. The intervention group participants were registered as end users of MyHealth Online and received no special services from the help desk. Support was limited to questions regarding the use and operation of the program. No self-management coaching was provided by any program resources. Intervention group participants received messages weekly via email reminding them to log in to MyHealth Online. Participants who fell below the desired threshold of participation (set at one log-in per week) received a message tailored to this condition, encouraging them to increase their participation and to contact the help desk if they required assistance to use the application. All control group participants received a message midway through the study reminding them to review the health education material and to contact the Carolinas HealthCare System help desk with any questions.

Statistical Analysis

Except for the Rasch psychometric statistics [41], all statistical analysis was performed using PASW statistics release 18 programs (SPSS Statistics GradPack; IBM Corporation, Somers, NY, USA). We used Winsteps version 3.6 (Rasch Measurement Software, Chicago, IL, USA) to calculate the Rasch person reliability and infit statistics for assessing the PAM-13's reliability and validity. Differences between the two groups were tested using the chi-square test of independence or Fisher exact test for categorical variables, and the *t* test for independent groups for pretest patient activation scores. These tests were also applied to assess between-group differences in the characteristics of patients who withdrew or dropped out of the study.

Interval-level data that were normally distributed resulted from this study, supporting the application of parametric models for the inferential tests. For testing the primary hypothesis, we evaluated the difference between groups in mean patient activation scores by applying analysis of covariance using the mean patient activation score at pretest as the covariate to reduce error variance. To test the second hypothesis and using a 2-step method, we analyzed change in patient activation scores between participants in the intervention group starting at different stages of patient activation. First, we divided intervention participants into three groups based on their stage of patient activation at the beginning of the study. We treated these groups as independent groups and applied 1-way analysis of variance to test for significance between groups. The 1-way analysis of variance is appropriate for comparing means between three groups but does not reveal the specific group differences underlying a significance difference [42]. To determine the specific groups (ie, baseline stage) demonstrating significant change, we conducted a post hoc test, using the Tukey honestly significant difference (HSD) method.

Results

Descriptive Characteristics

Table 1 presents demographic statistics for the control and intervention groups at the start of the study. The sample at baseline consisted of predominately non-Hispanic white persons between 45 and 64 years of age with a college degree. The sample consisted of slightly more women than men.

Table 2 shows health characteristics of the control and intervention groups at baseline. Most participants self-reported hypertension or diabetes as their main chronic condition.

Table 1. Demographic characteristics and between-group statistics at baseline.

Variable		Total sample (N = 201)		Control group (n = 100)		Intervention group (n = 101)		Test for difference	
		n	%	n	%	n	%	χ^2	*P* value
Age group (years)								$\chi^2_2 = 0.5$	.77
	25–44	36	18	16	16%	20	20%		
	45–54	66	33%	33	33	33	33%		
	55–64	99	49%	51	51%	48	47%		
Gender								$\chi^2_1 = 0.3$	.62
	Male	96	48%	46	46%	50	50%		
	Female	105	52%	54	54	51	50%		
Race								$\chi^2_1 = 0.6$	.43
	White	173	86%	88	88%	85	84%		
	Other	28	14%	12	12%	16	16%		
Ethnicity								$\chi^2_1 = 0.3$[a]	.58
	Hispanic, Latino, or Spanish origin	12	6%	5	5%	7	7%		
	Other	188	94%	94	95%	94	93%		
Education								$\chi^2_2 = 0.2$[a]	.90
	High school graduate or less	11	5%	5	5%	6	6%		
	Some college or trade school	66	33%	34	34%	32	32%		
	College graduate or more	123	62%	60	61%	63	62%		

[a] One participant assigned to the control group did not respond to this item.

Table 2. Health characteristics and between-group statistics at baseline.

Variable	Total sample (N = 201)		Control group (n = 100)		Intervention (n = 101)		Test for difference	
	n	%	n	%	n	%	χ^2	*P* value
Main chronic condition							$\chi^2_3 = 4.0$[a]	.27
Asthma	13	7%	7	7%	6	6%		
Diabetes	43	22%	16	16%	27	27%		
Hypertension	115	57%	60	60%	55	55%		
Other	29	14%	17	17%	12	12%		
Doctor visit in the past 6 months							$\chi^2_1 = 2.2$[b]	.13
Yes	139	70%	65	65%	74	75%		
No	60	30%	35	35%	25	25%		
Have used health education classes, support groups, or materials from doctor							$\chi^2_1 = 0.2$	.68
Yes	34	17%	18	18%	16	16%		
No	167	83%	82	82%	85	84%		
Reflecting on the past 6 months...I did different tasks and activities needed to manage my health condition so as to reduce my need to see a doctor							$\chi^2_1 = 0.0$	.94
Disagree	53	27%	26	26%	27	27%		
Agree	147	73%	73	74%	74	73%		

[a] One participant assigned to the intervention group did not respond to this item.

[b] Two participants assigned to the intervention group did not respond to this item.

At baseline no significant differences were found between the groups on any background variables, and the groups were not statistically different at the end of the study. Based on participants completing the study, there was no significant difference between the groups' mean pretest patient activation scores. Thus, the two groups at posttest were not significantly different from when the trial started. Furthermore, we found no significant difference between the groups in the average frequency of office visits during the study. This result suggests that any difference in patient activation scores between the groups at posttest was not influenced by patients seeing their doctors.

The control and intervention groups experienced attrition rates of 32% (32/100) and 41% (41/101) respectively, by the end of the study (Figure 1). We found no significance difference between the dropouts of the two groups in mean pretest patient activation scores ($t_{71} = .829$, $P = .41$). Furthermore, the average pretest score of dropouts was not significantly different from that of participants who completed the study ($t_{197} = –.951$, $P = .34$). These results indicate that participants who completed the study did not differ from those who did not in terms of the dependent variable at pretest.

Psychometric Properties of the PAM-13

The PAM-13 is based on the Rasch item response model [41]. Reliability of the PAM-13 is evaluated using the person reliability index; fit statistics are calculated to test construct validity of the measure [37]. The person reliability coefficient was a relatively high .83, showing a good spread of responses across the ordered items and expected endorsement patterns. Item and person infit statistics of .99 and 1.01, respectively, were well within the acceptable range [37]. In sum, the PAM-13 demonstrated strong psychometric properties in this study.

Effect of the Web-Based Intervention on Patient Activation

The Web-based intervention demonstrated a positive and significant effect on the patient activation levels, on average, of the participants in the intervention group. Controlling for the pretest patient activation scores, we found a significant difference in posttest patient activation scores between the two groups ($F_{1,123} = 4.438$, $P = .04$, effect size $r = .196$). Both groups experienced an average increase in patient activation scores during the study (Table 3), prompting an examination of change scores to determine whether the improvement in the control group may have attenuated the between-group difference at posttest. The difference in the mean patient activation score from pretest to posttest of the control group (mean 2.04, SD 10.01) was not significant ($t_{67} = 1.677$, $P = .10$), whereas the intervention group showed a highly significant change (mean 5.967, SD 9.70, $t_{57} = 4.683$, $P < .001$). The positive and significant change in the intervention group's patient activation reinforces the analysis of covariance results.

Table 3. Patient activation measure from pretest to posttest by group.

Group	n[a]	Patient activation score (pretest)		Patient activation score (posttest)		Analysis of covariance[b]	
		Mean	SD	Mean	SD	$F_{1,123}$	*P* value
Control	68	66.89	10.94	68.93	12.28	4.438	.04
Intervention	58	65.33	14.17	71.30	13.74		

[a] Group size (n) at posttest.

[b] Analysis of covariance on 12-week posttest scores controlling for patient activation score at pretest.

Baseline Patient Activation Stage and Change in Activation Scores

Prior to testing for differences in patient activation score changes between participants starting at each stage of patient activation, we combined patients at the first two stages into a single group to make this group's size comparable with the others (Table 4). The 1-way analysis of variance revealed that the mean change in patient activation scores across the three groups was significantly different ($F_{2,55}$ = 6.472, *P* = .003, effect size *r* = .436. Post hoc comparisons using the Tukey HSD test showed a significantly lower mean change score in the stage 4 group than in the combined stage 1–2 group (mean difference –8.457, 95% confidence interval –15.47 to –1.45; *P* = .01, effect size *d* = 0.45) and than among participants starting at stage 3 (mean difference –8.354, 95% confidence interval –15.06 to –1.64; *P* = .01, effect size *d* = 0.45). The difference in change scores between the stage 1–2 and stage 3 groups was nonsignificant (*P* = .999).

Despite beginning the study at the most advanced stage of patient activation, stage 4 group participants had the potential to substantially improve their patient activation scores. A ceiling effect was not evident in the stage 4 group, as 79% (22/28) of the participants began with activation scores between 68 and 80, sufficiently below the maximum score of 100, to experience a mean change in patient activation comparable with groups beginning at the earlier stages.

Table 4. Descriptive statistics for intervention group by patient activation stage at pretest.

Patient activation stage	n	Patient activation score (pretest)		Patient activation score (posttest)		Change in patient activation score	
		Mean	SD	Mean	SD	Mean	SD
1–2	14	46.11	4.42	56.19	9.50	10.08	7.87
3	16	61.38	3.85	71.35	11.35	9.98	10.64
4	28	77.2	7.72	78.82	10.32	1.62	8.26

Discussion

To our knowledge, this is the first study to examine the effect of a Web-based intervention designed to target various chronic conditions on patient activation, a construct encompassing a person's knowledge, skills, and self-efficacy to self-manage his or her health. It is also the first study to explore change in patient activation development stages among patients using a Web-based intervention. Results from this controlled trial suggest that a Web-based intervention targeting various chronic diseases and featuring the versatility to support self-management in the context of a person's particular chronic condition may improve the level of patient activation, a measure of self-management capabilities.

We found a small but significant difference in posttest patient activation scores between the control and intervention groups. The small effect size found between the groups' patient activation scores is consistent with other Web-based interventions demonstrating a significant effect on health behavior outcomes [24]. Within the intervention group, participants starting this 12-week study at the first three stages of patient activation improved their patient activation scores significantly compared with patients who started at the most advanced stage (4) of activation.

MyHealth Online is designed for use by patients with various health interests. Within this general-purpose architecture, applications are configurable to present content that targets a person's specific chronic condition. For this study, parameters were set to provide messages and health education modules to an intervention group participant based on the main chronic disease self-reported during the enrollment process. SMIs that embed general self-management skills training within interactive disease-specific modules are more effective than didactic disease education alone in enhancing the self-efficacy component of the patient activation construct [43]. Building on this evidence, health education modules in MyHealth Online contain interactive exercises that present a self-management problem in the context of a specific disease (eg, adhering to a hypertension medication regimen). This approach resembles the disease-specific education of Web-based interventions shown to enhance the self-efficacy of patients with arthritis [44] and diabetes [33], and is in contrast to a general-purpose Web-based intervention [35], which did not show significant improvement

in the self-efficacy component of patient activation. Thus, the present study suggests that, to improve patient activation, Web-based interventions designed for use by populations consisting of people with different chronic diseases should be designed to provide information therapy that is specific to a patient's medical condition.

People in the first three stages of patient activation are forming beliefs about their role in personal health and starting to build confidence in their abilities to self-manage [13]. From a chronic care management perspective, a desirable goal is to develop patients' self-management capabilities to the level reflected by the most advanced stage of patient activation. Patients with chronic conditions who are the most activated (ie, stage 4) are more likely to engage in self-management behaviors that are associated with improved clinical outcomes, including adhering to prescribed medication regimens, regularly testing glucose levels, and monitoring blood pressure [15]. The significant improvement in patient activation shown by intervention participants starting at the first three stages of patient activation suggests that MyHealth Online may have helped these patients become more capable of self-managing their health. Web-based interventions that aid the advancement of patients on the developmental continuum of patient activation contribute to an expanded population of patients who are able to actively manage their diseases—adhering to desirable health behaviors and knowing what steps to take when confronted with a condition-related incident (eg, diabetic hypoglycemia).

My Health Online appears to have aided primarily participants at the early stages of patient activation in gaining an understanding of their chronic conditions and developing the skills and self-confidence needed to better manage them. This apparent area of effectiveness of MyHealth Online is similar to findings from previous studies of the effects of Web-based interventions on the self-efficacy component of patient activation. Wangberg found that diabetic patients with the lowest baseline self-efficacy benefited the most from a Web-based intervention [45]. The nonsignificant change in stage 4 patients' activation scores in the current study is comparable with the observation by Berman et al that patients with high baseline levels of the more granular attribute of self-efficacy did not show significant change as a result of a Web-based intervention [46].

Highly activated (ie, stage 4) patients lack confidence in their abilities to solve problems they have not previously encountered [13]. Despite room for improvement among participants beginning at stage 4 of patient activation in this study, they did not appear to gain substantive benefit from the SMI that was used. This outcome suggests that the self-service and interactive health education modules of MyHealth Online did not improve the self-confidence of participants who possessed the knowledge and skills to self-manage their health in routine situations. They may require a more sophisticated SMI with scenarios for how to solve various complex problems not frequently encountered. More research is needed to understand the specific needs of patients at the most advanced stage of patient activation.

The contrast in change in patient activation of people at different starting points of self-management capabilities affirms the assertion by Hibbard and colleagues [13] that different SMIs are needed to help people on their journey along the continuum of patient activation. Although more research is needed, evidence suggests that a Web-based intervention should be designed to specifically target either people with low levels of patient activation or those in the advanced stage of patient activation. A Web-based SMI with features similar to those of MyHealth Online may yield the most optimal results when used by patients in the earlier stages of patient activation. By recognizing the different stages of patient activation and targeting Web-based interventions to patients in the earlier stages, chronic care program designers and practitioners may derive tangible value from this technology as measured by improvements in quality metrics for adherence and health status.

Strengths and Limitations

This study has several strengths and limitations. Important strengths were the rigorous randomized experimental design and posttest sample size (n = 128). The groups were statistically not different in terms of demographic characteristics and the dependent variable at baseline. Data met the assumptions for appropriate use of parametric tests. The dependent variable (patient activation score) was transformed to an interval-level scale, distributions of patient activation scores pre- and posttest were close to normal, variance of these scores shown by the two groups was sufficiently homogeneous, and the combined sample size exceeded the minimum needed to produce accurate results [42,47]. In addition, the PAM-13 demonstrated good psychometric properties.

Four limitations should be considered when interpreting the results. First, results may not be applicable to other populations and settings. Compared with the US population [4], the sample underrepresented people belonging to minority groups. Furthermore, 95% (189/200) of the study participants had attended college. This study follows a persistent pattern of research involving Web-based SMIs, which attract mostly college-educated, non-Hispanic, white participants [31,33,35,45,48,49]. Thus, the applicability of results from most research on Web-based interventions is inherently limited to a population willing and able to access the Internet. Further study is warranted to validate the results from this study with larger samples that are more representative of the US chronic disease population.

The second limitation is the unexplained reasons for the 41% attrition rate in the intervention group, a level higher than the 12%–25% reported for controlled trials involving Web-based interventions [44,46,50]. The six withdrawals and 24 dropouts may have been participants who experienced difficulties using the applications or had expectations for the Web-based intervention that were not met. These are problems reported by users of Web-based interventions in prior research [27,51]. The significance of usability or user expectations to this study's level of attrition is unknown, as measuring user satisfaction was not a study objective. Furthermore, although the two groups maintained statistical equivalency on all demographic and health descriptive variables at posttest despite the attrition, attributes of the dropouts not measured in this study may have influenced the study's outcomes.

Third, the scope of this research was limited to measuring change in patient activation when a Web-based intervention was used. The relative importance of specific functions and participants' perceived value of the Web-based intervention were not assessed. A fourth limitation is the possibility that participants were influenced by agents external to the study. Participants were encouraged to receive usual care; no restrictions were placed on use of other support resources.

Conclusion

Results from this study suggest that Web-based interventions have the potential to serve a vital role in health care providers' efforts to enhance the self-management capabilities of the growing population of chronically ill people. The study's outcomes reveal that the use of a Web-based intervention may result in more activated patients who are more proficient and self-confident in self-managing their chronic conditions. As a measure of self-management capabilities and a predictor of health outcomes, patient activation and its underlying components are mediators [52] in a causal chain of chronic care beginning with an SMI and ending with improved health outcomes. Activated patients are more likely to adhere to recommended health care practices, which in turn leads to improved health outcomes [12,53]. Thus, Web-based interventions that influence patient activation are an essential element of chronic disease management programs where success is measured in quality metrics such as patient adherence, functional status, or clinical effectiveness. Evaluating the effect of a Web-based intervention on patient activation provides program designers and practitioners with a more complete picture of the factors contributing to change in the health status of targeted patients.

As innovators work to transform health care to support the new "...era of chronic disease predominance" [11], they are encouraged to build on the discoveries from this and previous research to advance the industry's understanding of the value of Web-based SMIs in activating patients to take control of their health. By providing self-management support services on demand, anytime, and from anywhere, Web-based interventions have the potential to be a transformative technology in the management of chronic diseases, delivering self-management services to large numbers of people with fewer human resources than traditional SMIs have. To turn this potential into reality, evidence is needed that Web-based interventions can produce benefits for a sustained period among a diverse population. More clarity of the value of Web-based self-management tools as scalable interventions will stimulate the investments necessary to accelerate adoption throughout all segments of the large and growing chronic disease population.

Acknowledgments

No funding or sponsors supported this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6 [54].

[PDF File (Adobe PDF File), 810KB - jmir_v14i1e32_app1.pdf]

References

1. Bodenheimer T, Chen E, Bennett HD. Confronting the growing burden of chronic disease: can the U.S. health care workforce do the job? Health Aff (Millwood) 2009;28(1):64-74. [doi: 10.1377/hlthaff.28.1.64] [Medline: 19124856]
2. Pierson CA, Schumann L, Berner B. A perspective on chronic illness. J Am Acad Nurse Pract 2008 Mar;20(3):107-108. [doi: 10.1111/j.1745-7599.2007.00303.x] [Medline: 18336686]
3. Anderson G. Partnership for Solutions. 2004 Sep. Chronic Conditions: Making the Case for Ongoing Care URL: http://www.partnershipforsolutions.org/DMS/files/chronicbook2004.pdf [accessed 2012-02-06] [WebCite Cache ID 65G9pFwLw]
4. US Census Bureau. 2008. National Population Projections: Projections of the Population and Components of Change for the United States: 2010-2050 URL: http://www.census.gov/population/www/projections/summarytables.html [accessed 2011-07-18] [WebCite Cache ID 60H28O0Qv]
5. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. JAMA 2002 Oct 9;288(14):1775-1779. [Medline: 12365965]
6. Coleman K, Austin BT, Brach C, Wagner EH. Evidence on the chronic care model in the new millennium. Health Aff (Millwood) 2009;28(1):75-85. [doi: 10.1377/hlthaff.28.1.75] [Medline: 19124857]
7. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [Medline: 12867348]
8. Hill-Briggs F. Problem solving in diabetes self-management: a model of chronic illness self-management behavior. Ann Behav Med 2003;25(3):182-193. [Medline: 12763713]
9. Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. Lancet 2004;364(9444):1523-1537. [doi: 10.1016/S0140-6736(04)17277-2] [Medline: 15500899]

10. Hibbard JH. Moving toward a more patient-centered health care delivery system. Health Aff (Millwood) 2004;Suppl Variation:VAR133-VAR135. [doi: 10.1377/hlthaff.var.133] [Medline: 15471780]
11. Holman H, Lorig K. Patient self-management: a key to effectiveness and efficiency in care of chronic disease. Public Health Rep 2004;119(3):239-243. [doi: 10.1016/j.phr.2004.04.002] [Medline: 15158102]
12. Hibbard JH, Cunningham PJ. Center for Studying Health System Change. 2008. How Engaged are Consumers in Their Health and Health Care, and Why Does it Matter? HSC Research Brief no 8 URL: http://hschange.org/CONTENT/1019/ [accessed 2011-06-23] [WebCite Cache ID 5zet0wPpO]
13. Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Serv Res 2007 Aug;42(4):1443-1463. [doi: 10.1111/j.1475-6773.2006.00669.x] [Medline: 17610432]
14. Remmers C, Hibbard J, Mosen DM, Wagenfield M, Hoye RE, Jones C. Is patient activation associated with future health outcomes and healthcare utilization among patients with diabetes? J Ambul Care Manage 2009;32(4):320-327. [doi: 10.1097/JAC.0b013e3181ba6e77] [Medline: 19888008]
15. Hibbard JH, Tusler M. Assessing activation stage and employing a "next steps" approach to supporting patient self-management. J Ambul Care Manage 2007;30(1):2-8. [Medline: 17170632]
16. Bandura A. Self-efficacy: The Exercise of Control. New York, NY: WH Freeman; 1997.
17. Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. The extent and breadth of benefits from participating in chronic disease self-management courses: a national patient-reported outcomes survey. Patient Educ Couns 2007 Mar;65(3):351-360. [doi: 10.1016/j.pec.2006.08.016] [Medline: 17027221]
18. Mosen DM, Schmittdiel J, Hibbard J, Sobel D, Remmers C, Bellows J. Is patient activation associated with outcomes of care for adults with chronic conditions? J Ambul Care Manage 2007;30(1):21-29. [Medline: 17170635]
19. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026. [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
20. Warsi A, Wang PS, LaValley MP, Avorn J, Solomon DH. Self-management education programs in chronic disease: a systematic review and methodological critique of the literature. Arch Intern Med 2004;164(15):1641-1649 [FREE Full text] [doi: 10.1001/archinte.164.15.1641] [Medline: 15302634]
21. Marks R, Allegrante JP, Lorig K. A review and synthesis of research evidence for self-efficacy-enhancing interventions for reducing chronic disability: implications for health education practice (part II). Health Promot Pract 2005 Apr;6(2):148-156. [doi: 10.1177/1524839904266792] [Medline: 15855284]
22. 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]
23. Ritterband LM, Gonder-Frederick LA, Cox DJ, Clifton AD, West RW, Borowitz SM. Internet interventions: in review, in use, and into the future. Prof Psychol Res Pr 2003;34(5):527-534. [doi: 10.1037/0735-7028.34.5.527]
24. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
25. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
26. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008;15(6):729-736 [FREE Full text] [doi: 10.1197/jamia.M2547] [Medline: 18756002]
27. Kerr C, Murray E, Stevenson F, Gore C, Nazareth I. Internet interventions for long-term conditions: patient and caregiver quality criteria. J Med Internet Res 2006;8(3):e13 [FREE Full text] [doi: 10.2196/jmir.8.3.e13] [Medline: 16954123]
28. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
29. 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]
30. Solomon MR. Information technology to support self-management in chronic care: a systematic review. Dis Manag Health Outcomes 2008;16(6):391-401.
31. Jackson CL, Bolen S, Brancati FL, Batts-Turner ML, Gary TL. A systematic review of interactive computer-assisted technology in diabetes care. Interactive information technology in diabetes care. J Gen Intern Med 2006 Feb;21(2):105-110. [doi: 10.1111/j.1525-1497.2005.00310.x] [Medline: 16390512]
32. Cudney S, Sullivan T, Winters CA, Paul L, Oriet P. Chronically ill rural women: self-identified management problems and solutions. Chronic Illn 2005 Mar;1(1):49-60. [Medline: 17136933]
33. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VB, et al. Online diabetes self-management program: a randomized study. Diabetes Care 2010 Jun;33(6):1275-1281. [doi: 10.2337/dc09-2153] [Medline: 20299481]

34. Lorig KR, Ritter PL, Dost A, Plant K, Laurent DD, McNeil I. The Expert Patients Programme online, a 1-year study of an Internet-based self-management programme for people with long-term conditions. Chronic Illn 2008 Dec;4(4):247-256. [doi: 10.1177/1742395308098886] [Medline: 19091933]
35. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]
36. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin; 2002.
37. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930. [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
38. Insignia Health. Patient Activation Measure (PAM) 13 License Materials. Portland, OR: Insignia Health LLC; 2008.
39. Mettler M, Kemper DW. Information therapy: health education one person at a time. Health Promot Pract 2003 Jul;4(3):214-217. [Medline: 14610991]
40. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406 [FREE Full text] [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
41. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago, IL: University of Chicago Press; 1980.
42. Polit DF. Data Analysis & Statistics for Nursing Research. Stamford, CT: Appleton & Lange; 1996.
43. Francis CF, Feyer AM, Smith BJ. Implementing chronic disease self-management in community settings: lessons from Australian demonstration projects. Aust Health Rev 2007 Nov;31(4):499-509; discussion 498. [Medline: 17973606]
44. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008 Jul 15;59(7):1009-1017 [FREE Full text] [doi: 10.1002/art.23817] [Medline: 18576310]
45. Wangberg SC. An Internet-based diabetes self-care intervention tailored to self-efficacy. Health Educ Res 2008 Feb;23(1):170-179 [FREE Full text] [doi: 10.1093/her/cym014] [Medline: 17412717]
46. Berman RL, Iris MA, Bode R, Drengenberg C. The effectiveness of an online mind-body intervention for older adults with chronic pain. J Pain 2009 Jan;10(1):68-79. [doi: 10.1016/j.jpain.2008.07.006] [Medline: 18774342]
47. Norusis MJ. SPSS 16.0 Advanced Statistical Procedures Companion. Upper Saddle River, NJ: Prentice Hall; 2008.
48. McKay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. Diabetes Care 2001 Aug;24(8):1328-1334. [Medline: 11473065]
49. Ross SE, Moore LA, Earnest MA, Wittevrongel L, Lin CT. Providing a web-based online medical record with electronic communication capabilities to patients with congestive heart failure: randomized trial. J Med Internet Res 2004 May 14;6(2):e12 [FREE Full text] [doi: 10.2196/jmir.6.2.e12] [Medline: 15249261]
50. Glasgow RE, Nutting PA, King DK, Nelson CC, Cutter G, Gaglio B, et al. Randomized effectiveness trial of a computer-assisted intervention to improve diabetes care. Diabetes Care 2005 Jan;28(1):33-39. [Medline: 15616230]
51. Ralston JD, Revere D, Robins LS, Goldberg HI. Patients' experience with a diabetes support programme based on an interactive electronic medical record: qualitative study. BMJ 2004 May 15;328(7449):1159 [FREE Full text] [doi: 10.1136/bmj.328.7449.1159] [Medline: 15142919]
52. 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]
53. Bodenheimer T. Helping patients improve their health-related behaviors: what system changes do we need? Dis Manag 2005 Oct;8(5):319-330. [doi: 10.1089/dis.2005.8.319] [Medline: 16212517]
54. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

HSD: honestly significant difference
PAM-13: 13-item patient activation measure
SMI: self-management intervention

Edited by G Eysenbach; submitted 29.08.11; peer-reviewed by L Ritterband; comments to author 24.09.11; revised version received 24.01.12; accepted 27.01.12; published 21.02.12.

Please cite as:
Solomon M, Wagner SL, Goes J
Effects of a Web-Based Intervention for Adults With Chronic Conditions on Patient Activation: Online Randomized Controlled Trial
J Med Internet Res 2012;14(1):e32
URL: http://www.jmir.org/2012/1/e32/
doi:10.2196/jmir.1924
PMID:22353433

Original Paper

Social Influence as a Driver of Engagement in a Web-Based Health Intervention

Josée Poirier[1], PhD; Nathan K Cobb[2,3,4], MD

[1]MeYou Health, LLC, Boston, MA, United States

[2]Division of Pulmonary & Critical Care, Department of Medicine, Georgetown University Medical Center, Washington, DC, United States

[3]Department of Oncology, Georgetown University Medical Center/Lombardi Comprehensive Cancer Center, Washington, DC, United States

[4]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Josée Poirier, PhD
MeYou Health, LLC
1 Appleton Street, 4th Floor
Boston, MA, 02116
United States
Phone: 1 866 885 2822 ext 919
Fax: 1 617 728 3535
Email: josee.poirier@meyouhealth.com

Abstract

Background: Web-based health interventions can drive behavior change, but their effectiveness depends on participants' usage. A well-recognized challenge with these interventions is nonusage attrition or weak engagement that results in participants receiving low doses of the intervention, negatively affecting outcomes. We present an approach based on the theoretical concepts of social influence and complex contagion in an effort to address the engagement problem in a specific, commercial, online behavior change intervention.

Objective: To examine the relation between social ties and engagement within a specific online intervention. The aims were (1) to determine whether experiencing the intervention socially influences engagement, such that individuals with social ties show higher engagement than those without ties, and (2) to evaluate whether complex contagion increases engagement—that is, whether engagement increases as the number of ties an individual has in the intervention increases.

Methods: We analyzed observational data from 84,828 subscribed members of a specific Web-based intervention, Daily Challenge. We compiled three measures of engagement for every member: email opens, site visits, and challenge completions (response to action prompts). We compared members with and without social ties within the intervention on each measure separately using 2-tailed independent-sample *t* tests. Finally, we performed linear regressions with each simple engagement measure as the dependent variable and number of social ties as the independent variable.

Results: Compared with those without social ties, participants with social ties opened more emails (33.0% vs 27.2%, $P < .001$), visited the website more often (12.6 vs 6.7 visits, $P < .001$), and reported completing more of the actions they were prompted to perform (11.0 vs 6.1 actions, $P < .001$). Social ties were significant predictors of email opens (beta = 0.68, $P < .001$), site visits (beta = 1.52, $P < .001$), and reported action completions (beta = 1.32, $P < .001$).

Conclusions: Our initial findings are higher engagement in participants with social ties in the program and are consistent with the view that social influence can drive engagement in a Web-based health intervention.

(J Med Internet Res 2012;14(1):e36) doi:10.2196/jmir.1957

KEYWORDS

Web-based health interventions; engagement; social networks; social influence

Introduction

Although Web-based health interventions can drive behavior change across multiple conditions [1], their effectiveness highly depends on participants' usage. Adherence remains a well-recognized challenge with Web-based interventions, with rapidly decaying amount of exposure to program content (nonusage attrition; [2]) presenting a significant barrier to the

development of effective systems. For instance, many Web-based interventions reported in the literature are affected by rapid attrition [2-4] and suboptimal consumption of program content, as measured by site visits and time spent on the site [5-7]. Despite this, such systems have also demonstrated a dose response, with increasing levels of adherence associated with improved outcomes [8-11]. Capitalizing on the potential of these systems to reach large proportions of the population will require solving the twin problems of adherence and ongoing consumer engagement.

Social influence may offer a solution to problematically weak engagement and adherence to an intervention as designed. The term refers to the ability of connected individuals to affect one another's thoughts, ideas, and behaviors. Social influence contributes to the spread of behaviors through social contact [12,13], such that individuals adopt a new behavior more readily if their social ties display it. This effect may reflect individuals' reliance on the actions of others to determine the appropriate behavior in a given situation (informational social influence), or their underlying desire to conform to the expectations of others (normative social influence) [14]. Thus, in a health intervention, facilitating interactions between participants and exposing them to one another's activity may induce high behavior-adoption rates. When applied to program usage behavior, this approach could harness social influence to improve engagement in Web-based interventions and ultimately affect behavioral change outcomes. The social approach has been successful in previous Web-based health interventions (eg, QuitNet; [15]) and may be the most promising in terms of cost effectiveness and success rate [16].

In addition to their potential scalability at a minimal cost, social Web-based interventions can magnify the potential effect of social influence by markedly increasing social contact. The phenomenon of complex contagion has been reported across various domains such as technology, fashion, migration, urban legends, and, more recently, health [17,18]. It refers to the observation that individuals who receive social reinforcement from multiple contacts tend to adopt behaviors in higher numbers and do so more rapidly [17]. In social Web-based interventions, social reinforcement can occur particularly frequently. The multiplicity of ties a participant has increases the probability that he or she will be repeatedly exposed to others engaging in a given behavior. Hence, the very nature of social interventions offers a favorable environment for complex contagion to facilitate the adoption of a desired behavior such as high program usage.

In this study, we describe one real-world implementation in which theoretical concepts of social influence and complex contagion are applied to address the engagement problem. We present evaluation results of this approach, used to drive program usage in a specific, commercial, online health and wellness intervention.

This intervention, named Daily Challenge and delivered primarily on the Web and by email, allows members to form connections with other members, effectively building an individual social network within the intervention. The network provides various forms of social influence that may increase engagement: social proof, accountability, and support. The members' activity is visible to their social network, and vice versa. Therefore, one can observe a contact (a connection or friend in the system) displaying a behavior (such as returning to the site, forming a new friendship, trying a new feature, or reporting having completed a challenge) within the intervention and adopt it as well (social proof). Further, the transparency of activity over every tie renders individuals accountable to every person in their network, potentially altering behavior. The visibility of members' activity in the intervention enables their ties to detect and act on others' struggles or inactivity by providing support and direct, personalized nudges (formalized communications, similar to a "wink" or "poke" in other products). Similarly, individuals can offer positive support to their connections in the form of encouragement, companionship, and information. We hypothesized that this combination of social proof, social accountability, and social support would increase engagement in the program.

The goal of this study was to explore the relation between social ties and engagement within this specific online intervention. Our aims were 2-fold: (1) to determine whether experiencing the intervention socially influences engagement, such that individuals with social ties show higher engagement than those without ties, and (2) to evaluate whether complex contagion increases engagement—that is, whether engagement increases as the number of ties an individual has in the intervention increases.

Methods

Intervention Description

Daily Challenge [19] is a publicly available health and wellness intervention designed to assist individuals in making small, health-related, positive changes to improve their overall well-being. The intervention, which undergoes continual development, is based on work demonstrating that the adoption of small behavioral changes across the larger population can have a significant public health benefit [20,21]. Individuals can register through Facebook. Each day after registration they receive, most commonly via email, a suggestion of a small action they can realistically accomplish that day. For example, an individual might be challenged to take the stairs at work, or to review the salt content in today's lunch. Information on how to complete the day's action and how it relates to well-being accompanies the suggestion, which is presented as a challenge for the day. Members can report that they have completed the challenge by clicking a Done button via a link in the email or on the website directly (Figure 1, top left). Members who do not acknowledge completion by 4 PM local time receive an optional reminder (Figure 1, top right).

Once they indicate they have completed the challenge, members are prompted to share how they did it (Figure 1, top left), either with everyone in the local (Daily Challenge) community, or solely with their Daily Challenge connections. These social connections are formed within Daily Challenge. To establish a connection, a member sends a friendship request to the member with whom he or she would like to connect. The other member may be an extant social tie (eg, a family member, friend,

colleague, or Facebook friend) or someone the member met through Daily Challenge and might have interacted with in the public threads. If the request is accepted, the connection is formed. All connections require confirmation from the other party and are thus reciprocal.

Members with connections have access to the primary social features in the program: they can view the private and public "how I did it" notes of their personal connections (Figure 1, top left); they can send encouragement to a personal connection; or they can enter a pact to complete challenges together for 5 consecutive days. Additionally, connected members can follow each others' activity more closely. Members can easily see when their connections complete challenges—among other actions—in an activity stream (Figure 1, bottom). The reminder emails also detail which of the members' ties have completed the day's challenge and how they did it (Figure 1, top right). Thus, members do not need to return to the site to be exposed to social proof.

As Daily Challenge is an operating, real-world intervention, maximizing participation rates is paramount. As such, minimal demographics beyond those provided automatically via Facebook are gathered. Various game mechanics are used in addition to social factors to encourage participation, including the awarding of points and badges.

Figure 1. Screenshots of the Web-based intervention Daily Challenge. Top left: A challenge with its Done button, as seen on the site. Top right: An email reminder with social proof. Bottom: Activity stream in which members can follow their connections' activity, including challenge completions.

Participants

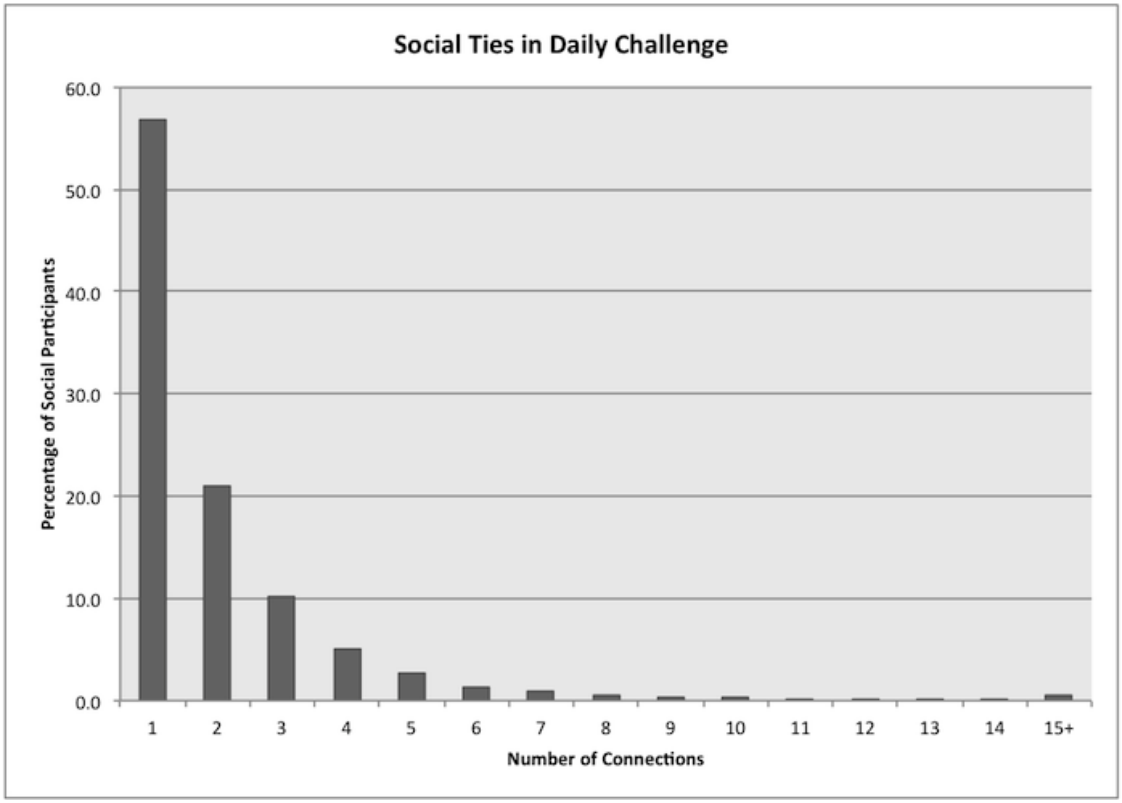
Participants were Daily Challenge members who registered between September 14, 2010 and July 20, 2011 and who had been members for at least 30 days (N = 128,233). We examined the first 30 days of activity in the program. Individuals whose only interaction with the system was registration were excluded (inactive: n = 43,405 or 33.8%). Members with at least minimal interaction with the system—defined as at least one known email open or site visit after their first emailed challenge (during the 30-day window)—were included in the analysis (active: n = 84,828 or 66.2%). The final dataset comprised 84,828 members who had formed a total of 67,648 personal connections within Daily Challenge. Table 1 summarizes the attributes of the participants in the study. Figure 2 shows the distribution of ties in the program for social users.

The data analyzed in this study are stored in real time in a relational database used for quality improvement purposes. Data include available demographics, self-reported challenge completions, and observed metrics of utilization. All data were anonymized prior to extraction and analysis.

Table 1. Attributes of the 84,828 participants in the study.

Attribute	n (%) or mean (SD)	
Days in program, mean (SD)	119.1	(61.6)
Personal connections, mean (SD)	0.8	(1.8)
Gender, n (%)		
Women	71,031	83.7%
Men	8,712	10.3%
Unknown	5,085	6.0%

Figure 2. Distribution of the number of connections social participants formed within Daily Challenge.



Process Measures

We obtained engagement measures at the individual level for the following three actions: (1) an email open, (2) a visit to the website, and (3) the completion of a challenge. Participants can perform each of these actions with or without social ties in Daily Challenge.

Email Opens

We marked an email as opened if the participant's email client downloaded the embedded graphical images in the message. This metric undercounts actual reading of the messages, as some email programs either are text based or do not load images. Not every participant subscribed to email delivery. For those who did, they further varied in whether they requested morning emails (a 7 AM message delivering the day's challenge), reminder emails (sent out at 4 PM if the participant had not yet

completed the challenge), or both. Because of the optional and dynamic nature of this feature, we defined email opens as a percentage of days on which a member opened at least one email that was sent to them.

Site Visits

Site visit refers to the total number of times a member visited the website. Users are uniquely tracked by identifiers embedded in URLs in email messages, long-term cookies, or, if needed, via a login screen. We recorded a new visit if a page was seen after a period of at least 30 minutes without participant-generated activity while the participant was logged in. Consequently, the maximum number of visits was 47 a day.

Challenge Completions

Completions is the number of challenges *reported* by the member as completed. Participants who click the Done link from an email are automatically recorded as having completed the challenge. Alternatively, participants can click the Done button directly on the website to report their challenge completion. Within the context of this study we make no assumptions about the user's actual behavior in the real world, but rather use the behavior of reporting a completion as a marker of active engagement with the system.

The variables personal connections, email opens, site visits, and completions showed a Pareto distribution, failing to meet the assumption of normality, and were subject to a log-normal transformation prior to analysis. We conducted analyses on transformed individual-level data with the statistical software R [22].

Results

As shown in Figure 3, participants formed the bulk of their ties (n = 50,834 or 75.1%) within 7 days, the overall majority on signup day (n = 36,230 or 53.6%).

We began our investigation of the relation between social ties and engagement by contrasting social and nonsocial experiences of the program. Participants were divided into two groups: participants with at least one personal connection in Daily Challenge (social participants) and those without (nonsocial participants). These groups were compared on each measure of engagement (Table 2 for conditional means).

Figure 3. Distribution over time of friendship formation for participants' first month in Daily Challenge. X-axis: Participants' days in the program. Y-axis: Percentage of connections formed.

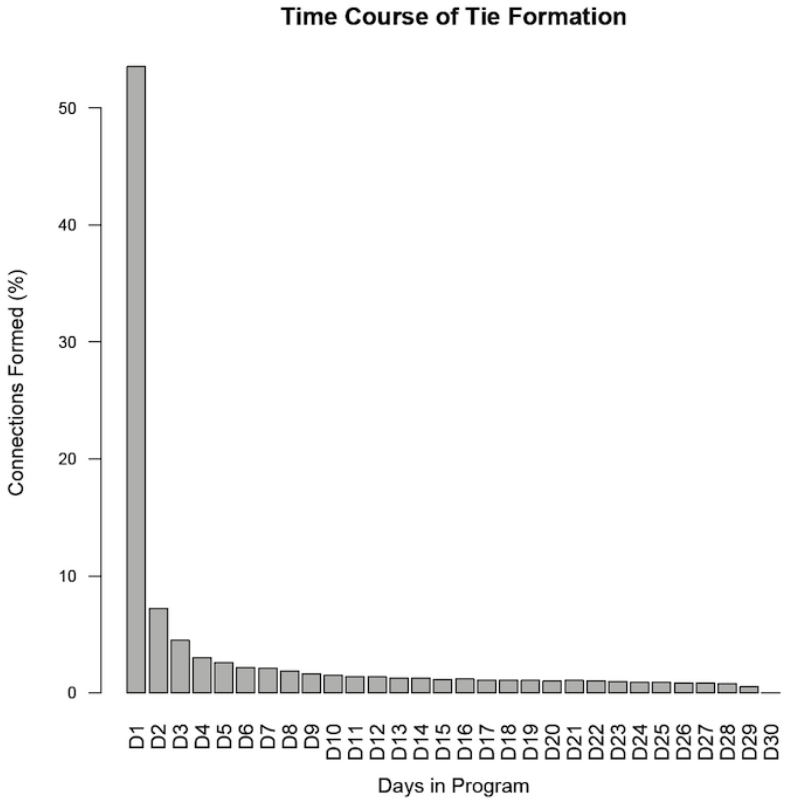


Table 2. Untransformed conditional means for nonsocial and social participants.

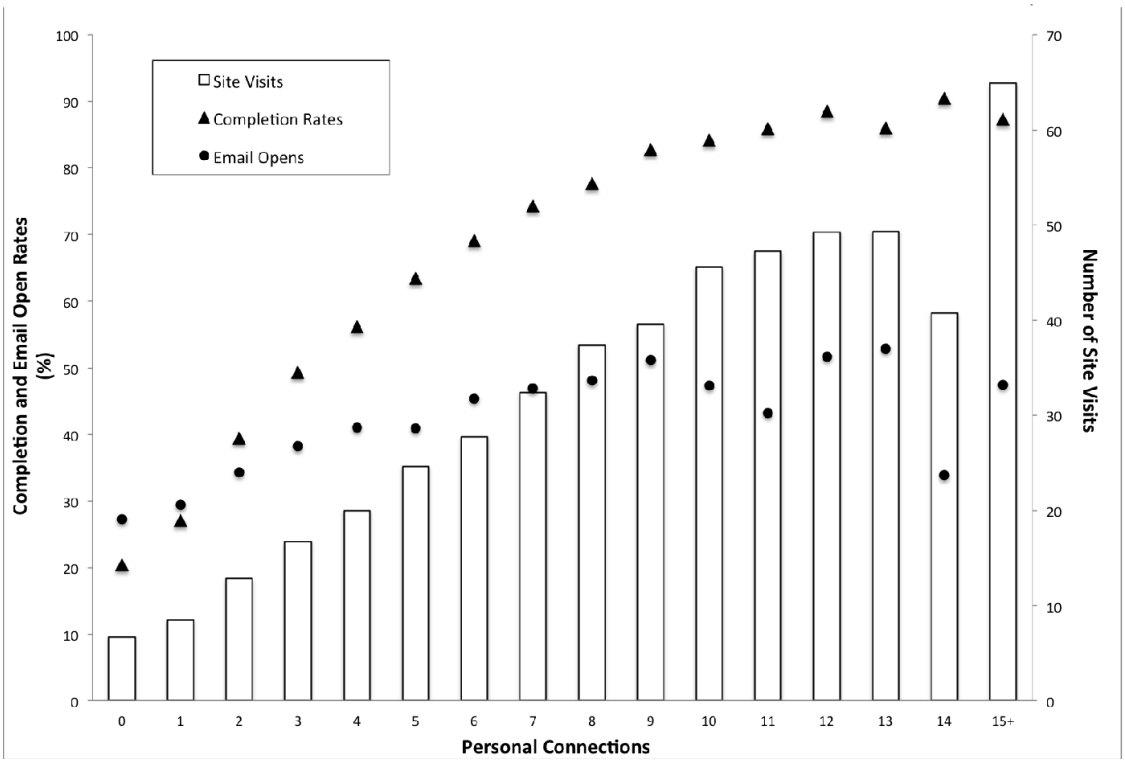
	Nonsocial	Social
Email opens (%)	27.2%	33.0%
	(n = 51,775)	(n = 33,021)
Site visits (total)	6.7	12.6
	(n = 51,787)	(n = 33,041)
Completions (total)	6.1	11.0
	(n = 51,787)	(n = 33,041)

Across the board, participants with social ties in the program showed higher engagement. Specifically, participants with personal connections opened more emails sent to them than did participants who had not formed social ties ($t_{65,198}$ = 24.2, P < .001, 2-tailed; 95% confidence interval [CI] 5.3−6.2). Furthermore, social participants visited the website more often than their nonsocial counterparts ($t_{51,767}$ = 61.4, P < .001, 2-tailed; 95% CI 5.7−6.0). The pattern was the same with respect to challenge completions, where social participants showed higher rates than nonsocial participants ($t_{57,749}$ = 75.7, P < .001, 2-tailed; 95% CI 4.7−5.0).

We conducted a series of linear regressions to substantiate the relation between social ties and engagement. Figure 4 presents aggregate data on the effect of social ties on each single engagement metric. Predictive models with number of personal connections as the predictor of engagement explained variance in email opens (adjusted $R2$ = .0094), site visits (adjusted $R2$ = .1436), and completions (adjusted $R2$ = .1324).

A positive effect of social ties was found for each measure of engagement. Number of personal connections in the program related to email open rates (beta = 0.68, P < .001), site visits (beta = 1.52, P < .001), and challenge completions (beta = 1.32, P < .001).

Figure 4. Email opens, site visits, and completion rates per number of personal connections in Daily Challenge. For illustrative clarity, aggregate data are shown. The category 15+ encompasses less than 0.2% of the participants.



Discussion

In this study, we explored the relation between social ties and engagement in a specific online health behavior intervention. We aimed to determine whether experiencing the intervention socially influenced engagement and whether complex contagion increased engagement. We found compelling evidence that a social experience of the program was associated with improved adherence and program retention, and that engagement level was higher in participants with more social ties in the program.

This social effect affected all measures of engagement examined: email opens, site visits, and reported challenge completions. Participants interacted with the intervention more often if they had social ties, despite the fact that having social ties does not affect a member's ability to perform the actions studied (or the reward received for each action). The

improvement in engagement was substantial: for instance, social participants reported completing an average of 11.0 challenges over 30 days, whereas nonsocial participants reported 6.1 challenge completions for the same period. These findings suggest that social interventions are more engaging than nonsocial ones, and that building social interventions may be part of a solution to the problem of engagement and adherence.

To enable participants to be social, interventions need not require complex, large networks, as a single tie seems to suffice to bolster engagement. However, the social effect was more pronounced for participants with larger social networks, although its benefits seemed to diminish beyond 12 connections, suggesting a natural threshold. More specifically, a higher number of ties was associated with more frequent email opens, site visits, and challenge completions. These results are consistent with the phenomenon of complex contagion, whereby an individual is more likely to adopt a behavior observed in multiple other individuals. By building exposure to others' behavior into the intervention, we aim to create an environment that enables complex contagion.

Our findings provide early evidence that interventions exploiting social influence can extend to natural, participant-created networks. Online social interventions can leverage participants' extant networks while providing an environment in which individuals can effectively extend and strengthen their networks by adding new connections. Thus, larger social networks supply the redundancy to facilitate behavior adoption (complex contagion) and the wealth of resources of an enriched network, resources that can satisfy needs otherwise unmet.

Together, our results support the hypothesis that social relationships can affect engagement in an intervention. We note that it is possible that self-reported challenge completions either did not represent true offline behavior or, moreover, had no lasting impact. However, challenge completions, in addition to email opens and site visits, are objective indicators of engagement and dose delivered: participants demonstrated engagement in the action of clicking the Done button. Thus, our findings objectively substantiate the benefits of integrating a social network system on engagement and dose delivered within a health behavior intervention.

There are limitations to this early exploration. The data used were gathered for quality improvement purposes and lack many useful descriptors of participants, including full demographics or indicators of health status. Given that the majority of ties are formed shortly after registration, we chose to treat individual's social networks as static and the data as cross-sectional; it is possible that the evolution of an individual's network over time has a different effect on behavior. It also remains unknown to what degree ties can be augmented or encouraged within such a system—in other words, whether individuals' propensity to form beneficial connections is fixed or mutable.

Most important, it could be the case that individuals more inclined to be adherent initially form social ties, and that these ties are a marker, and not a driver, of adherence and compliance. However, the influence of social ties on one's attitude and behavior is well documented in social psychology (eg, [14,23,24]) and social network science (eg, [17,25]). The resistance to adopting a risky, costly, or controversial behavior can be overcome by positive reinforcement from multiple independent sources (complex contagion; [17]). Within this model, the continuous exposure of participants to social messages defining active participation as normative leads to an increase in participation proportional to the number of contacts. Our findings are consistent with both this model and highly controlled translational trials in online settings [18] that demonstrate similar findings.

These exciting but initial findings raise several questions on the relation between social ties and engagement. First, without a randomized design, it is difficult to conclusively demonstrate directionality. Importantly, the exact mechanisms of the social effect remain to be determined. For instance, it is unknown whether accountability to one's ties, social proof of a tie's activity, or social support (among others) influences engagement. Careful examination of the nature and frequency of social interactions between ties may help tease apart the relative contributions of these potential mechanisms.

The moderators of the social effect should also be examined. The influence of a social tie may be affected by the attributes of the connected individuals or of the connection itself. In particular, social influence may be strongest between participants with shared attributes, or who interact in a specific manner. Moderators may also affect social influence above the individual level, at the community or larger network levels. Furthermore, the topology of an individual's network, or their position in the network, may modulate social influence. Modern social network methods provide tools to explore these factors and their impact.

Ultimately, these data add weight to the emerging evidence that health behavior change systems structured to include social features can deliver a greater dose of the intervention over time. The careful inclusion of these features, coupled with early and ongoing evaluation, has the potential to augment engagement, retention, and adherence to an intervention, and ultimately drive behavior change.

Acknowledgments

The authors acknowledge Stephen Sopp for his assistance with the data. This research was supported by MeYou Health, a subsidiary of Healthways, Inc.

Conflicts of Interest

Dr Poirier is Director of Program Design and Research for MeYou Health. Dr Cobb serves as Science Advisor to MeYou Health, developer of Daily Challenge. Both authors have a personal financial interest in MeYou Health.

References

1. 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]
2. 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]
3. 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]
4. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer 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]
5. Leslie E, Marshall AL, Owen N, Bauman A. Engagement and retention of participants in a physical activity website. Prev Med 2005 Jan;40(1):54-59. [doi: 10.1016/j.ypmed.2004.05.002] [Medline: 15530581]
6. 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]
7. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]
8. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216. [doi: 10.1080/14622200500055319] [Medline: 16036277]
9. Heesch KC, Mâsse LC, Dunn AL, Frankowski RF, Mullen PD. Does adherence to a lifestyle physical activity intervention predict changes in physical activity? J Behav Med 2003 Aug;26(4):333-348. [Medline: 12921007]
10. Jacobs N, Clays E, De Bacquer D, De Backer G, Dendale P, Thijs H, et al. Effect of a tailored behavior change program on a composite lifestyle change score: a randomized controlled trial. Health Educ Res 2011 Oct;26(5):886-895. [doi: 10.1093/her/cyr046] [Medline: 21712501]
11. 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] [Medline: 18984557]
12. Rogers EM. Diffusion of Innovations. New York, NY: Free Press; 1995.
13. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007 Jul 26;357(4):370-379 [FREE Full text] [doi: 10.1056/NEJMsa066082] [Medline: 17652652]
14. Cialdini R, Trost M. Social influence: social norms, conformity, and compliance. In: Gilbert DT, Fiske ST, Lindzey G, editors. The Handbook of Social Psychology. 4th edition. Boston, MA: McGraw-Hill; 1998:151-192.
15. 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] [Medline: 20466971]
16. Christakis NA. Health care in a web. BMJ 2008 Jun 28;336(7659):1468. [doi: 10.1136/bmj.a452] [Medline: 18583678]
17. Centola D, Macy M. Complex contagion and the weakness of long ties. Am J Sociol 2007;113(3):702-734. [doi: 10.1086/521848]
18. Centola D. The spread of behavior in an online social network experiment. Science 2010 Sep 3;329(5996):1194-1197. [doi: 10.1126/science.1185231] [Medline: 20813952]
19. MeYou Health LLC. 2010. Daily Challenge URL: https://challenge.meyouhealth.com/signup/ [accessed 2012-02-20] [WebCite Cache ID 65b0wHE29]
20. Hill JO, Wyatt HR, Reed GW, Peters JC. Obesity and the environment: where do we go from here? Science 2003 Feb 7;299(5608):853-855. [doi: 10.1126/science.1079857] [Medline: 12574618]
21. Hill JO. Can a small-changes approach help address the obesity epidemic? A report of the Joint Task Force of the American Society for Nutrition, Institute of Food Technologists, and International Food Information Council. Am J Clin Nutr 2009 Feb;89(2):477-484 [FREE Full text] [doi: 10.3945/ajcn.2008.26566] [Medline: 19088151]
22. R Development Core Team. R: A Language and Environment for Statistical Computing: Version 2.13.1 (2011). Vienna: R Foundation for Statistical Computing; 2008.
23. Kelman HC. Compliance, identification, and internalization three processes of attitude change. J Conflict Resolut 1958;2(1):51-60. [doi: 10.1177/002200275800200106]
24. Latane B. The psychology of social impact. Am Psychol 1981;36(4):343-356. [doi: 10.1037/0003-066X.36.4.343]
25. Valente TW. Social network thresholds in the diffusion of innovations. Soc Networks 1996;18(1):69-89. [doi: 10.1016/0378-8733(95)00256-1]

Abbreviations

CI: confidence interval

Edited by G Eysenbach; submitted 03.10.11; peer-reviewed by M Wanner, B Cugelman; comments to author 03.11.11; revised version received 29.12.11; accepted 16.01.12; published 22.02.12.

Please cite as:
Poirier J, Cobb NK
Social Influence as a Driver of Engagement in a Web-Based Health Intervention
J Med Internet Res 2012;14(1):e36
URL: http://www.jmir.org/2012/1/e36/
doi:10.2196/jmir.1957
PMID:22356829

Review

Short Message Service (SMS) Applications for Disease Prevention in Developing Countries

Carole Déglise[1,2], MIH, MD; L. Suzanne Suggs[3], MS, PhD; Peter Odermatt[1,2], MPH, PhD

[1]Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

[3]Institute of Public Communication, Faculty of Communcation Sciences, Università della Svizzera italiana, Lugano, Switzerland

Corresponding Author:
Carole Déglise, MIH, MD
Department of Epidemiology and Public Health
Swiss Tropical and Public Health Institute
Socinstrasse 57, Postfach
Basel, 4002
Switzerland
Phone: 41 61 284 8111
Fax: 41 61 284 8101
Email: cardeglise@yahoo.com

Abstract

Background: The last decade has witnessed unprecedented growth in the number of mobile phones in the developing world, thus linking millions of previously unconnected people. The ubiquity of mobile phones, which allow for short message service (SMS), provides new and innovative opportunities for disease prevention efforts.

Objective: The aim of this review was to describe the characteristics and outcomes of SMS interventions for disease prevention in developing countries and provide recommendations for future work.

Methods: A systematic search of peer-reviewed and gray literature was performed for papers published in English, French, and German before May 2011 that describe SMS applications for disease prevention in developing countries.

Results: A total of 34 SMS applications were described, among which 5 had findings of an evaluation reported. The majority of SMS applications were pilot projects in various levels of sophistication; nearly all came from gray literature sources. Many applications were initiated by the project with modes of intervention varying between one-way or two-way communication, with or without incentives, and with educative games. Evaluated interventions were well accepted by the beneficiaries. The primary barriers identified were language, timing of messages, mobile network fluctuations, lack of financial incentives, data privacy, and mobile phone turnover.

Conclusion: This review illustrates that while many SMS applications for disease prevention exist, few have been evaluated. The dearth of peer-reviewed studies and the limited evidence found in this systematic review highlight the need for high-quality efficacy studies examining behavioral, social, and economic outcomes of SMS applications and mobile phone interventions aimed to promote health in developing country contexts.

(J Med Internet Res 2012;14(1):e3) doi:10.2196/jmir.1823

KEYWORDS

Mobile health; developing countries; prevention; SMS; short message service

Introduction

The last decade has seen an unprecedented growth of the development of information and communication technologies (ICT) infrastructure worldwide [1]. The trend exhibits a clear shift from landline toward mobile phones, whose subscription rates are three times greater than for landlines. The expansion has been most significant in the developing regions, where diffusion reached more than 40% of the population in 2007, thus connecting millions of previously unconnected people. Developing countries not only have the majority of world mobile phone subscribers, but will also account for 80% of the new ones [2].

Mobile telephones offer short message service (SMS), also known as text messaging. SMS is a communication protocol standardized in the Global System for Mobile communications allowing messages of 160 characters maximum to be interchanged from a mobile phone or a computer to one or many mobile phones simultaneously [3]. SMS can send information in near-real time to thousands of people as recipients of standardized, bulk messages or even personalized or tailored messages. SMS is available on all cellular phones, including cheap low-end handsets, through the Global System for Mobile communications network.

Recent research has focused largely on the use of SMS for health purposes in developed countries, while compelling studies seem to remain anecdotal in the developing world [4-13]. But text messages could have considerable implications for disease prevention efforts in developing countries where their potential is recognized not only in their communication attributes, such as voice call and text messages, but also in their data transfer capabilities. In addition, the robustness, ease of maintenance, and relative affordability of handsets, compared with computers, make the mobile phone and SMS very attractive in resource-poor areas with little electricity and slow Internet connections. As such, an increasing number of mobile health initiatives are being implemented and tested in the developing world [14].

Primary and secondary prevention encompass all actions that help in "averting the occurrence of disease" and "halting the progression of a disease from its early unrecognized stage to a more severe one," respectively [15]. As declared in the Ottawa Charter, health promotion is "the process of enabling people to increase control over, and to improve, their health. It supports personal and social development through providing information, education for health, and enhancing life skills" [16]. Today, the World Health Organization recognizes that "advanced information and telecommunication technologies should be employed to their fullest extent wherever possible, in order to create effective and transparent communication channels that will allow interactive sharing and learning among various groups of stakeholders in the society" [17].

With the rapid expansion of mobile health applications, combined with current and predicted economic and public health challenges, reviews of existing applications and evidence of SMS-supported interventions in developing countries are needed. Therefore, the objective of this study was to examine SMS-supported interventions for prevention of communicable and noncommunicable diseases in developing countries. We began with a review of peer-reviewed and gray literature on existing SMS-based interventions. Gray literature is "that which is produced on all levels of government, academics, business, and industry in print and electronic formats but which is not controlled by commercial publishers." In other words, "grey literature includes documents that have not been formally published in a peer-reviewed format" [18]. Then, we assessed the effectiveness of SMS-based interventions and identified drivers and inhibitors to adoption. Finally, we present recommendations for future research and practice.

Methods

The data used in this study are part of a larger review aiming to examine SMS-supported interventions for prevention, surveillance, management, and treatment compliance of communicable and noncommunicable diseases in developing countries. We used a subject-based approach to systematically search both peer-reviewed and gray literature published in English, French, and German up to May 2011. The decision to include gray literature was motivated by the fact that public health professionals often use gray literature to inform decision making at the practice, program, and policy levels [18], and that both gray and scholarly evidence should be included in reviews of new and innovative topics in order to reduce bias and provide good estimates of the effects of interventions [19].

Data Sources and Search Strategy

Bibliographic database and search engines were PubMed (incorporating Medline), EMBASE, CINAHL, PsycINFO, ScienceDirect, SpringerLink, EBSCOhost, ABI/INFORM, Google Scholar, and Google. Synthesis-producing organizations comprised the Cochrane Collaboration, the Centre for Reviews and Dissemination, and the Campbell Collaboration.

The search terms were *health promotion*, *public health*, *preventive health services*, *disease prevention*, *population surveillance*, *patient compliance*, *patient care management*, *disease management*, *self care*, *HIV*, *acquired immunodeficiency syndrome*, *tuberculosis*, *communicable diseases*, *communicable disease control*, *chronic disease*, *noncommunicable disease*, *cellular phone*, *mobile phone*, *cell phone*, *SMS*, *MMS*, *text messag**, *picture messag**, *developing countries*, *developing world*, *Asia*, *Africa*, *South America*, *Latin America*, and *Europe*, *Eastern*. We used the medical subject headings in Medline and Emtree tools in EMBASE. When necessary, these words were translated into French and German to minimize language bias.

The gray literature search strategy was inspired by a guideline for finding gray literature in public health [18]. Requests for cases and papers were posted on email lists, through personal connections, and directly to health organizations and SMS providers. Key gray literature sources were the Digital Repository Infrastructure Vision for European Research, the Electronic Theses Online System, the Networked Digital Library of Theses and Dissertations, the Directory of Open Access Repositories, DissOnline.de, SUDOC Catalogue, Social Science Research Network, Scirus, and Catholic Media Council Library. Other sources of gray literature comprised proceedings from related conferences, funding and not-for-profit organizations, listservs, blogs, SMS, and mobile phone providers. For identified projects, we contacted the corresponding agency to request documents and white papers that described the application and evaluation results. References of included documents and citation tracking of articles were also searched.

Inclusion and Exclusion Criteria

Following a detailed protocol, an initial reviewer identified and retrieved all eligible documents. The titles and abstracts were first examined to remove obviously irrelevant reports. A second reviewer confirmed eligibility and relevance. The full text of

the potentially relevant ones was reviewed to finally select eligible papers.

English-, French-, and German-language documents were included when they were related to interventions using SMS in developing countries for disease prevention. For the purpose of this review, SMS referred to all applications that used SMS or texting for disease prevention purposes. These included applications that automatically shifted between SMS and general packet radio service channels, had live-person components, and connected to the Internet. Developing countries were defined as *developing and emerging economies*, a World Economic Outlook classification system based on (1) per capita income level, (2) export diversification, and (3) degree of integration into the global financial system [20]. For this study, we considered countries classified as developing and emerging economies in the International Monetary Fund's World Economic Outlook report in 2008 to be developing countries (eg, Algeria, Brazil, Cambodia, India, Mexico, Poland, and Zambia) [20].

Documents were excluded if the applications did not use SMS, if they were not used in developing countries, or when they did not focus on the prevention of communicable and noncommunicable diseases. Documents published in languages other than English, French, or German were also excluded.

Data Collection and Management

Eligible articles were reviewed and data extraction was inspired by a template provided in the *Cochrane Handbook for Systematic Reviews of Interventions* [21]. Applications were given a unique identification number. Disease focus, country or countries of implementation, project objectives, target audience, status (planned, ongoing, or completed), funding source, and contact details were recorded. When an evaluation was reported, the methods used and outcomes measured (health outcome, process of care, relevance, acceptability, and cost effectiveness) were recorded. Each personal communication attempt was documented along with contact information and the outcome of the effort. Finally, SMS-based interventions used for health promotion and disease prevention were extracted from the larger dataset and used for this review.

Results

The first screening identified 4008 citations. We retrieved 38 additional papers by personal communications (n = 21) and by hand searching the reference lists of eligible articles (n = 17). After an initial screening of abstracts and titles, we excluded most because they were related to developed countries or examined harmful effects of mobile phone use. A second selection was done to exclude applications that did not focus on disease prevention (eg, applications for surveillance, disease management, or compliance with a treatment).

We identified 34 different SMS-based prevention applications according to the eligibility criteria (Table 1 [22-67]), among which 5 included details about an evaluation (Table 2 [27,36,41,49-51,63]). Most were from gray literature sources. The interventions addressed a variety of topics, with human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) being the most common (n = 18), followed by sexual and reproductive health (SRH) (5). Some applications addressed multiple diseases (Table 3). Although projects extended across regions, SMS initiatives were concentrated in South Africa, Kenya, and India (Figure 1).

Table 1. Short message service (SMS) interventions for disease prevention

Intervention (reference)	Country	Disease	Description	Comments
It begins with you [22]	30 African countries	HIV/AIDS[a]	SMS voting system on what happened on the show, asked viewers to share what they have done to advance an HIV-free generation, and encouraged all Africans to start by knowing their HIV status.	Target: population; status: ended after 1 season
Star Project [23]	6 African countries	HIV/AIDS	Counterpart to India's *Freedom from HIV* project. SMS used for downloading 2 mobile phone games (AIDS Fighter Pilot and AIDS Penalty Shoot Out) to raise HIV/AIDS awareness, deployed on low-end and sophisticated colored devices.	Target: population; technology: ZMQ; specificity: developed English and 2 local languages (Kiswahili and Shen)
Talk Back [24]	Botswana	HIV/AIDS	Weekly television program for HIV prevention, broadcasted live, to stimulate interactivity with teachers and viewers through phone lines, SMS, emails, and letters.	Target: teachers and students
UNICEF[b] [25]	Central African Republic	Measles, malaria, diarrhea	Multimedia campaign used SMS to encourage vaccination, use of long-lasting insecticidal nets, and handwashing.	Target: parents of young children
Text Me! Flash Me! [26,27]	Ghana	HIV/AIDS	Health education and promotion messages sent to mobile phone numbers collected by peer educators and social networks. Clients who text in "HELP" were referred to live helpline counselors, who called back within 24 hours.	Target group: most-at-risk populations: men who have sex with men and female sex workers
eQuest [28]	Kenya	HIV/AIDS	Contest engaged youth in discussions about HIV/AIDS. Youth sent SMS answers to questions about HIV/AIDS received on their mobile phone, after checking information in a special eQuest column printed in the newspaper.	Target: youth; incentives: airtime, T-shirts, mobile phone, computers, DVD players, and a home theatre system
Makutano Junction [29,30]	Kenya	HIV/AIDS	Soap opera based in a fictitious Kenyan village supported by SMS. Viewers were invited to text in if they needed more information on a given topic.	Target: population
Mobile4Good [31]	Kenya	HIV/AIDS	"My question" allowed customers to anonymously ask HIV/AIDS and breast cancer-related questions and receive answers via SMS. "Health Tips" provided subscribers with useful tips on various pertinent health issues via SMS.	Target: population

Intervention (reference)	Country	Disease	Description	Comments
Afriafya [32]	Kenya	HIV/AIDS	Community resource centers worked with information and communication technology to access various information, including on health, via SMS request or other means of communication. Answer was sent back by email, booklet, or SMS.	Target: rural population; technology: telecenter
Pariah News [33]	Madagascar	HIV/AIDS	Citizen media-enabled project that broadcasted HIV/AIDS message via SMS, Internet radio, and blogs.	Target: sex workers; technology: Ushahidi platform, open source
Health On Line [34]	Mali	HIV/AIDS, malaria	Social marketing campaign that used bimonthly free SMS with health slogans and reference to an interactive sexual health website.	Target: young, urban people (n = 350,000)
Learning about Living [35,36]	Nigeria	HIV/AIDS and SRH[c]	Health promotion and prevention was based on HIV/AIDS, SRH, maternal morbidity, and gender violence with (1) MyQuestion: HIV/AIDS-related questions sent by public via SMS, Web, or hotline, answered by trained counselors, (2) MyAnswer: prizes won by texting correct answer to a quiz.	Target: young people; incentive: airtime; scaleup: in existing and new states
RapidSMS [37]	Nigeria	Malaria	SMS helped deploy bed nets by (1) tracking commodities from state stores to distribution points by monitoring coupon distribution, (2) sending SMS reminders about distribution times and location for beneficiaries.	Target: population; technology: RapidSMS (UNICEF innovation); license: open source
Beat It [38,39]	South Africa	HIV/AIDS	Free SMS to enter the draw for prizes that motivated people to check results on *Beat It* television program. Designed to promote positive living, treatment access, and HIV infection prevention.	Target: youth; technology: Cell-Life; incentives: mobile phone, airtime
Cell-Life [40,41]	South Africa	HIV/AIDS	Mass messaging for prevention, linking clinic and patients to peer-to-peer support and counseling at no charge, through a computerized capture of mobile phone number and automatic SMS back with the information.	Target: patient; technology: Cell-Life; license: open source; multicomponent project; status: ongoing

Intervention (reference)	Country	Disease	Description	Comments
Project Masiluleke [42,43]	South Africa	HIV/AIDS	Project provided several mobile phone-based applications for HIV/AIDS care: "Access Information" and "Get Tested". Health promotion messages broadcasted in unused space of "Please Call Me," a free form of SMS widely used in Africa.	Target: population; technology: SocialTxt from Praekelt Foundation; license: open source; multicomponent project compliance
South African Depression and Anxiety Group [44]	South Africa	Mental health	National toll-free suicide helpline and SMS for adolescents in crisis.	Target: young people; status: ongoing
Digital mosquito net vouchers [45]	Tanzania	Malaria	Implemented long-lasting insecticidal net distribution using SMS voucher system for controlling counterfeited voucher.	Target: pregnant women
Kimasomaso [46,47]	6 African countries	SRH	Radio program transmitted voices of young people keeping audio diaries, associated with helpline. Also provided SMS to redirect callers and text senders to local support.	Target: young people
AppLab [48]	Uganda	HIV/AIDS, SRH	Leveraged existing Village Shared Phone Operators to deliver mobile information services in health and agriculture with (1) SMS-based health tips and searchable database, (2) "Clinic Finder", to locate nearby health clinics and services.	Target: population; technology: AppLab applications
Text to Change [49-51]	Uganda	HIV/AIDS	Interactive SMS quiz designed to help resolve key issues around HIV transmission and prevention, in the form of a multiple choice questionnaire that guaranteed free voluntary counseling and testing services to participants who answered correctly. Three quizzes offered weekly in English.	Target: population (15,000); incentives: voluntary counseling and testing services, airtime and mobile phone; status: ongoing, plan for Uganda and other African countries
UNICEF [52]	Zimbabwe	Cholera	Nationwide SMS information campaign during larger cholera campaign.	Target population
China Netcom [53]	China	SRH	SRH education and awareness campaign with SMS and hotline that gave access to medical experts.	Target: population and teenagers
SARS[d] education [54,55]	China	SARS	Mobile phone subscribers could call an SMS that alerted them if they were within 1 km of a SARS-infected building, where confirmed cases existed, and about news updates.	Target: population; license: proprietary, mobile operator

Intervention (reference)	Country	Disease	Description	Comments
Indonesia: Community Based Avian Influenza Control Project [56]	Indonesia	Avian influenza	SMS-based contest to encourage travelers in buses to be careful and to test their knowledge on the diseases.	Target: population; incentives: airtime
Condom Condom Campaign [57,58]	India	HIV/AIDS	Condom use promotion and HIV/AIDS awareness campaign among young men with (1) SMS opinion to vote on HIV/AIDS issues, (2) condom-themed mobile phone ringtone using SMS to get a push in reply, from where the user could download the ringtone.	Target: men; incentives: mobile phone and free talk time
Freedom HIV/AIDS [59]	India	HIV/AIDS	SMS used for (1) downloading mobile phone games to raise HIV/AIDS awareness, deployed on low-end and sophisticated colored devices, (2) announcement of radio shows on HIV/AIDS, (3) information on the nearest HIV testing center.	Target: population; technology: ZMQ; specificity: developed in local languages
Heroes Project [60,61]	India	HIV/AIDS	Multiple media channels including SMS to get key messages on HIV/AIDS out to the general public.	Target: population
Indian tuberculosis campaign [62]	India	Tuberculosis	Public awareness campaigns used SMS for tuberculosis information.	Target: population
Breast cancer awareness [63]	India	Cancer	SMS as reminder to conduct breast self-examination.	Target: working women in private companies
Global Hand-washing Day/UNICEF [64]	Nepal	Diarrhea	Public awareness campaign used SMS to encourage hand-washing.	Target: population
Mobilink [65]	Pakistan	Polio	Broadcasted millions of SMSs to encourage parents to get their children vaccinated against polio.	Target: parents; specificities: initiative of services provider
Sex-Ed Text [66]	Philippines	SRH	Computerized system using SMS to receive and then return the keyword of interest for getting complete and free information.	Target: young people
CardioNet [67]	Mexico	Cardiac diseases	Public prevention campaign in which users took a quick cardiac assessment screening by SMS.	Target: population; technology: Voxiva; license: proprietary

[a]Human Immunodeficiency Virus/acquired Immunodeficiency Syndrome.

[b] United Nations Children's Fund.

[c] Sexual and Reproductive Health.

[d] Severe acute respiratory syndrome.

Table 2. Short message service (SMS) interventions for disease prevention with an evaluation

Intervention	Disease	Aims	Methods	Results	Limitations/challenges
Learning About Living [36]	HIV/AIDS[a] and SRH[b]	To document and distribute lessons learned during the process of initiating, planning, implementing, and monitoring the project.	Project evaluation; duration: 14 months after implementation; N = 9000 youth. Outcomes: objective and subjective.	User description: median age 24 years, 93% from urban and suburban settings, 79% male, 60,000 questions received by SMS, multiple use of service 49%. User satisfaction: 76%, 24% free, 12% prompt response, 7% easy access and availability, 24% educative HIV and SRH. Reason for dissatisfaction: >50% no answer due to bad network, 25% question partially answered, 18% response time too long, 7% question misunderstood.	Limitations: methods not described, results not comparable from state to state due to partner variation in evaluation methods. Challenges: network fluctuations, spam messages, girls and rural outreach.
Cell Phone for Life [41]	HIV/AIDS	To evaluate how the pilot service is perceived by the organizations running the service, those receiving the messages and those close to the recipients who may be affected: (1) baseline study of expectations of the pilot service, current access to information and service usage, and behavior, (2) process evaluation to assess how the pilot service is received, (3) review outcomes and results from the pilot service.	Baseline survey; N = 210. Outcomes: objective and subjective.	Participant characteristics: mean age 26.32 years, 71% female, 66.2% own a mobile phone, 79.5% comfortable using mobile phone, 64.3% use SMS. Qualitative information: Important: two-way communications important, messages may be consulted at any point, save time and money. Concerns: maintenance cost to members, the poor have no access to phones with required technology to perform Cell-Life functions technology, people change mobile phone regularly, potential for misuse or private use of SMS allocated by Cell-Life.	Challenges: maintenance and SMS costs, limited access for poor people, high mobile phone turnover, potential misuse or private use of SMS by Cell-Life.

Intervention	Disease	Aims	Methods	Results	Limitations/challenges
Text Me! Flash Me! [27]	HIV/AIDS	To understand the main reasons for contacting the helpline.	Pilot project; duration: 6 months; N = 1169 calls; 12 health workers and 135 MSM[c] randomly selected. Outcome: objective and subjective, interview and focus groups.	Participants: reach in first month: 5 counselors counseled 439 MSM clients = average of 88 MSM clients per counselor per month compared with 50 MSM clients per peer educator or health worker per month in facilities and communities; 87% shared the information with others: 40% did it by SMS to a mean of 8.6 persons; increased knowledge and intention to use condoms; 47% went to health services as counseled voluntary counseling and testing uptake increased after launch of campaign.	Challenges: lack of monetary incentives related by counselors.
Text to Change [49,50]	HIV/AIDS	Understand satisfaction with and use of pilot system.	Pilot project; duration: 8 weeks; N = 15,000. Outcome: objective, number of answers, and voluntary counseling and testing attendance.	Participants: 2610 actively participated, 807 texted back their age (mean 29.2 years), 801 texted back their gender (70.8% male; 29.2% female). Number of questions/participant: >1 (17.4%); voluntary counseling and testing attendance: 255 people.	Challenges: language barriers, confidentiality concerns (33.8%), lack of information on voluntary counseling and testing participants.
Text to Change [51]	HIV/AIDS	Assess access to and interest in receiving health information on mobile phone.	Survey; duration: 6 months. Outcome: number of answers.	Participants: 1506 with response rate of 86.7%; 62% male, 42% between 12 and 14 years and 51% between 15 and 17 years, 27% owned a mobile phone, among whom 93% sent SMS over past 12 months (34% every day, 35% weekly, 21% monthly, 9% <1/month), 19% of those who sent SMS said they did it to get health information in the last year; 51% of all adolescents said they were somewhat or extremely likely to access health education program through SMS, which was associated with owning a mobile phone; high-risk adolescents were equally likely to be interested in receiving HIV prevention program via SMS.	Limitations: low rate of mobile phone ownership in Mbarra, self-administered questionnaire, with results relevant to secondary school students with good English literacy. Large number of missing data.

Intervention	Disease	Aims	Methods	Results	Limitations/challenges
Breast cancer awareness [63]	Cancer	To assess SMS effectiveness as reminders for making women aware of breast cancer.	Participants: 106 women; duration: 6 months; SMS to remind to conduct breast self-examination.	Participants: 20–54 years old. Outcome: among those who forgot, 54% forgot and intended to do it, 47% were busy, and 4% had some questions regarding the exam. After 2 months of reminders, the practice of breast self-examination improved significantly.	Limitations: small sample, little information on methods. Women working in private sector.

[a] Human immunodeficiency virus/acquired immunodeficiency syndrome.

[b] Sexual and reproductive health.

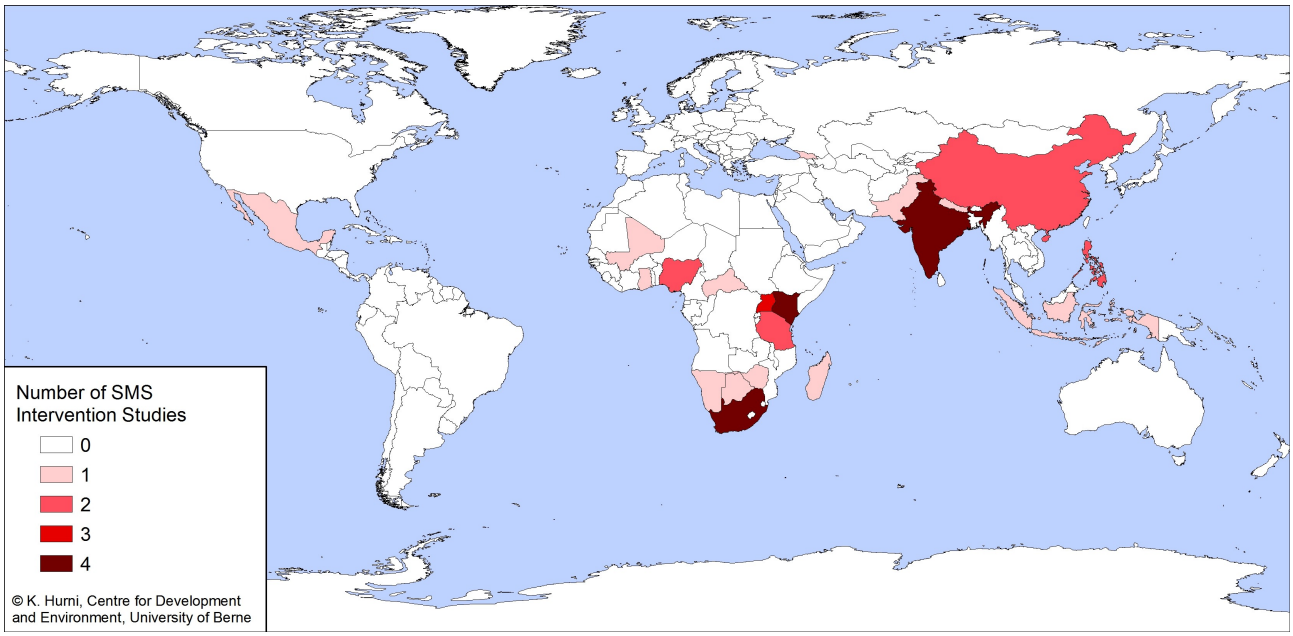
[c] Men who have sex with men.

Table 3. Short message service (SMS) for disease prevention: disease focus of 34 applications

Disease focus	Number of applications	Percentage of sample
Human immunodeficiency virus/acquired immunodeficiency syndrome	18	47%
Sexual and reproductive health	5	13%
Malaria	4	11%
Diarrhea	2	5%
Others[a]	9	24%

[a] Includes Measles, Avian Influenza, Cholera, Severe Acute Respiratory Syndrome, Tuberculosis, Poliomyelitis, Mental Health, Cancer, and Cardiac Disease. Some Applications Related to Several Diseases.

Figure 1. World map showing the distribution of short message service (SMS) intervention studies in developing countries (number of studies per country).



Description of the Applications

The 34 applications, listed in Table 1 [22-67], were described using varying levels of detail about the purpose of the initiative and features. Most provided information about the method used for communication (one-way or two-way), the use of games and contests, and incentives to increase the adoption and use of the application.

One-way Communication

SMSs served as one-way communication tools for prevention. SMSs were sent outbound to large numbers of subscribers who

had no opportunity to respond to messages or seek specific advice. The communication was standardized or targeted, tended to aim at a large population, and provided information about healthy behaviors and testing services. Such initiatives often required the participation of mobile phone operators who provided a database of phone numbers used during the campaign. As such, SMSs were sent with health promotion slogans for HIV/AIDS in the Heroes Project in India [60,61], to encourage parents to get their children vaccinated during Polio Vaccination Days in the Mobilink project in Pakistan [65], by the United Nations Children's Fund for cholera information during an outbreak in Zimbabwe [52], to prevent diarrhea by encouraging hand-washing in Nepal [64], and to provide information about tuberculosis in India [62]. They were also sent to parents of young children encouraging vaccination, bed net use, and hand-washing in the Central African Republic [20]. For malaria prevention, text messages reminded the population about bed net distribution dates, times, and locations [37].

Cell Phone for Life in South Africa offered an open source system to support the fight against HIV/AIDS [40,41]. One component of the project diffused awareness through one-way mass messaging. The first stage of Project Masiluleke [42,43] circulated HIV/AIDS and tuberculosis messages to the general public. SMSs were sometimes accompanied by links to interactive website with, for example, information on sexual health for the project Health On Line in Mali [34]. In one component of Text Me! Flash Me!, SMSs were sent to subscribers with educational and promotional messages, either for a general communication strategy or in response to trends noticed through ongoing quality monitoring and evaluation [26,27].

Two-way Communication

Two-way communication interventions included opportunities for people to text in for health tips, obtain tailored information about clinic locations, or contact a live person. SMS also supported other media channels allowing people to ask questions when they needed more information. They could remain anonymous, which is particularly important for stigmatizing issues such as HIV/AIDS, other SRH issues, and tuberculosis.

In the project Mobile for Good, "Health Tips" were sent to Kenyan subscribers who asked questions on HIV/AIDS and breast cancer-related issues [31]. An automatic computerized system was also established in the Philippines. In the Sex-Ed Text project, people who texted "SET" to a specific number (ie, short code) received a menu of keywords for SRH. When they responded with the keyword of interest they received free information [66]. Some projects leveraged existing public access services to deliver the intervention. Launched in Uganda, the Application Laboratory used shared phone operators to deliver information through an SMS-searchable database of health tips on sexually transmitted diseases, family planning, and maternal health, as well as details on local clinics and timetable of operation for outreach services [48]. In China, individuals could text in to receive the locations of buildings affected by severe acute respiratory syndrome and know of cases within 1 km of the calling location [54,55]. Users could also text in for daily updates on the syndrome. Messages were in both English and Chinese.

Several initiatives used television and radio programs in combination with SMS to engage in two-way communication. In the television soap opera *Makutano Junction*, viewers were encouraged to send an SMS if they needed more information about HIV/AIDS [29,30]. During the first, year, 30,000 text messages were received, and audience research showed that such a medium was useful to deliver information to rural and periurban audiences. Similarly, the It Begins With You campaign used SMS to engage young people in watching a show and voting with SMS [22]. The campaign aimed to get young people to share what they are doing to advance an AIDS-free generation and to know their HIV status. The Botswana Talk Back initiative aimed to provide education about HIV and involved a television talk show where viewers were encouraged to SMS with the presenters in the studio during live broadcasts [24]. *Kimasomaso* was a radio series that provided education about SRH issues [46,47]. Individuals could send in a text and receive support and advice.

Another component of the previously mentioned Cell-Life project used two-way communication to provide information on the location of voluntary counseling and testing (VCT) clinics [40,41]. Anybody could dial a specific phone number associated with a certain piece of information. A computer captured the number and sent an SMS back with the requested information at no charge to the requester. In Kenya, Afriafya provided community resource centers with technology access to HIV information [32]. Individuals requested information through SMS, and information was returned by SMS, email, or in print. Pariah News in Madagascar used SMS to send news about HIV/AIDS testing and prevention to sex workers [33]. Sex workers updated their health and HIV status and reported cases of violence anonymously, willingly, and in real time.

The Project Masiluleke connected mobile users to existing call centers, where trained operators provided health information, counseling, and referrals to local testing clinics [42]. After 3 weeks, average daily call volume to the national AIDS helpline tripled in Johannesburg, South Africa. Another component of the HIV/AIDS education initiative Text Me! Flash Me! in Ghana allowed for two-way communication. Targeting men who have sex with men and sex workers, users could text HELP to the project shortcode and a live peer counselor called back within 24 hours [26]. The South African Depression and Anxiety Group provided a national toll-free suicide crisis line and an SMS for adolescents in crisis [44]. In Tanzania, pregnant women could receive digital vouchers for bed nets, to prevent malaria [45]. Doctors sent an SMS to a retail bed net outlet and received back a valid digital voucher code. That code was transferred to the woman, who could then pick up a bed net. This also reduced the spread of fraudulent paper vouchers. In China, subscribers of two local mobile providers could text with experts about SRH issues [53].

Another project evaluated the use of SMS on breast cancer prevention, targeting working women in the private sector in Delhi. An SMS was sent every month according to the person's last menstrual cycle to remind her to conduct breast

self-examination at the end of her menstrual period. The content not only reminded the women to conduct breast self-examination, but also asked them to send an SMS back stating whether they had done it [63].

Games and Contests

The Indian Freedom HIV/AIDS project [59] and its African counterpart, the Star Programme [23], both used games to increase participation and promote disease prevention behaviors (Figure 2 [59]). Both used mobile platforms that were deployable on low-end black-and-white as well as sophisticated color cell phones to purposely target different preferences and socioeconomic groups. In India, 10 million game sessions were played over a 15-month period, and in Africa 6 million games were played in the first year. The project later included other topics such as tuberculosis, malaria, swine flu, lifestyle, and women's health issues.

The Condom Condom Campaign in India used different approaches to encourage young men to communicate about HIV/AIDS and to use condoms [57,58]. An SMS opinion vote was published in print media around the core message "smart men talk about condoms." In six newspapers across four states, an average of 89% of readers responded to a text vote. In addition, to get people to talk about condoms, the project developed a condom-themed ringtone that people could download via SMS shortcodes (ie, dedicated 5-digit mobile phone numbers). There were 270,000 requests for it.

Other issues addressed by games and contests included avian flu and heart disease. In 2005, the Community Based Avian Influenza Control Project in Indonesia, an SMS-based contest, encouraged travelers in buses to test their knowledge of Avian Influenza [56]. In Mexico, CardioNet used a quick cardiovascular risk assessment screening by SMS [67]. In turn, the numbers collected were used to send advice and information about heart health. The company was also developing a comparable tool for breast cancer information.

Figure 2. Freedom HIV/AIDS game: Quiz with Babu. Reproduced with permission [59].

Incentives

Some projects generated participation with a rewards system. Text To Change in Uganda designed an SMS multiple choice quiz, three times per week, to raise HIV/AIDS awareness and increase VCT [45]. Parallel to radio and newspaper advertisements, background information was provided in a special column in a magazine. Users submitting correct answers to the quiz were offered free VCT services and the possibility to win mobile phones or airtime. The program also monitored knowledge and beliefs through responses. The project Beat It in South Africa used free SMS to enter a prize draw where the results were announced during a television show designed to promote positive living and HIV/AIDS prevention [38,39]. This drew viewers and provided a contactable database. Campaigns such as Learning about Living in Nigeria (Figure 3) [35,36,68] and eQuest in Kenya [28] often targeted youth, encouraging them to correctly answer questions on HIV/AIDS issues for free airtime, T-shirts, mobile phones, computers, or DVD players.

Figure 3. Learning about Living. Reproduced with permission [35,68].

Evaluations and Outcomes

Most of the reported applications did not provide information about an evaluation. Five interventions reported outcomes, with varying levels of detail (Table 2 [26,27,36,41,49-51,63]): Learning About Living [36], Cell for Life [41], Text Me! Flash Me! [27], Text to Change [49-51], and Breast Cancer Awareness project [63]. Each of these is described below including available information about the aim and method used to evaluate it. Findings are described, and stated benefits and limitations are then noted.

The project Learning about Living, an HIV and SRH education campaign, was evaluated after 14 months of implementation [36]. The project reached 9000 young people (median age of 24 years) who submitted 60,000 HIV/AIDS-related questions by SMS, and 4500 people (50% of users) accessed the service more than once. Most users (93%) were from urban or suburban setting and were male (79%). The majority (76%) were satisfied with the service, mostly because it was free (24%), quick (12%), or easily available (7%). Among dissatisfied users (24%), half did not receive a response to their SMS due to mobile network fluctuation, and 50% complained about the timing of answers. Despite the lack of information on methodology and variation in methods of evaluation between different sites, this initiative

may have increased access to information for young people. The strengths of the project were regular formal and informal communication between all stakeholders. Noted challenges included network fluctuations, spam issues, and low reach of women and rural populations.

The pilot service of Cell-Life was evaluated using quantitative information from 210 respondents as well as qualitative inputs from 10 in-depth interviews and 5 focus groups [41]. Among the 210 participants, 66% owned a mobile phone, 80% felt comfortable using a mobile phone, and 64% were comfortable using SMS. Awareness of Cell-Life was poor, and its primary role was perceived as information dissemination at multiple levels. Users ranked the two-way communication and 24-hour availability of the messages for consultation on the mobile phone as important aspects, with the principle advantage that Cell-Life would save them time and money. Many requests were received for information about HIV and health in general, medication reminders, clinic locators, and access to reference sources of experts. Barriers included concerns about maintenance and SMS costs, limited access for poor people, high mobile phone turnover, and potential misuse or private use of SMS.

The Text Me! Flash Me! Service was an HIV/AIDS education initiative in Ghana that strived to strengthen HIV/AIDS knowledge among the most-at-risk population. After 6 months of implementation, the initiative was evaluated [27]. The aim was to understand the main reasons for using the service. Participants reported sharing information they received with others (87%) and 40% of them did so by SMS. The most popular service of the system was referral to VCT services (71% of the calls), which resulted in a substantial increase in VCT uptakes. Traditionally, counseling occurred at a rate of 50 clients per week per counselor. This program resulted in each counselor providing advice for approximately 88 clients per week. Reported challenges included a lack of financial incentives for counselors.

The project Text to Change aimed to increase the use of VCT centers. The system was evaluated in several papers, including 1 published in peer-reviewed literature [49-51]. In 2 papers, the project showed a 40% increase of VCT uptake after using an SMS quiz [49,50]. Identified challenges included language barriers due to English-language messages, a majority of male participants, confidentiality concerns (33.8%), and missing data (no information collected at VCT centers). The low number of respondents highlighted the importance of marketing at the beginning of the campaign and reassuring anonymity. In the quantitative study published in 2011 [51], the reported response rate was higher (86.7%) and still characterized by a majority of male participants. Among the SMS users, 19% mentioned they used it to get health information, and slightly more than half (51%) said they would be extremely likely to access health education programs through SMS, which was correlated with owning a mobile phone. High-risk and low-risk adolescents were equally likely to be interested in receiving HIV prevention programs via SMS. This study suggested that not all adolescents would have access to a mobile phone intervention and that nearly half would not necessarily choose to receive HIV/AIDS health information via this channel. It shows that adolescents may engage with education in different ways, thus emphasizing the importance of making HIV/AIDS prevention messages available in different modes and environments.

In India, the community-based study to evaluate the effect of SMS on breast cancer prevention targeting working women in the private sector in Delhi reported a significant increase of women performing breast self-examination after 2 months of SMS reminders [63]. The study sample was small and targeted only highly educated women working in the private sector who own mobile phones.

Discussion

This review collected descriptive information about SMS-based disease prevention interventions in developing countries, reported as of May 2011. Applications were widespread with small clusters in South Africa, Kenya, and India. Most of them targeted HIV/AIDS and SRH, but the topics were broad. Many SMS-delivered applications were initiated by projects with modes of intervention varying between one-way and two-way communication. Some used educative games or incentives to increase participation.

Of the 34 applications, 5 reported an evaluation. The evaluations primarily focused on the processes and outputs of the interventions. However, some provided evidence on behavior change outcomes, such as increase in uptake of VCT services and increase in breast self-examination. Available data suggest that for the 5 initiatives evaluated, SMS interventions were feasible, were well accepted by the targeted population, and allowed for high reach. Additionally, several projects that did not conduct or report an evaluation did report usage statistics that suggest the target audience used the service (at varying levels of engagement).

Some issues appeared as potential barriers to usage. These included lack of timely responses, mobile network connection fluctuations, lack of financial incentives, maintenance and SMS costs, high mobile phone turnover, and potential misuse or private use of SMS. Language barriers were identified as limiting factors in one campaign [49,50]. While some countries such as South Africa have 11 official languages and many more indigenous ones [69], scalability of the intervention may depend on the ability to develop health promotion messages in various languages and ethnocultural contexts, and finding ways to target people speaking less-known languages. A key learning of the consultation was that ICTs are “only as good as the information they seek to communicate” and the importance of considering “the whole communication process of which ICTs have opened up unprecedented opportunities.” [5] It is important to develop culturally and gender-sensitive messages and approaches provided in the languages of the target population. Transmitting messages through intermediaries, developing a voice recognition system, or using pictograms or graphs may be worthwhile solutions to language barriers. The limited length of SMS (160 characters) also influences content; however, next-generation mobile services will allow the transmission of richer messages and multimedia capacities, affording new opportunities for more effective communication.

There were also concerns about maintenance and SMS costs, high mobile phone turnover, and reliability and security of systems. Reliability may be questioned by the sender, who does not know whether the messages have been delivered, and by the receiver, who may not receive the message [70]. This was of particular importance with the high mobile phone turnover reported [41]. Users raised concerns about the ability to ensure data protection and confidentiality [49,50], even though the mobile phone was perceived as the most trustworthy technology over others [70]. These aspects were particularly important in settings where users shared mobile phones. It has been an undeniable concern for patients with tuberculosis and HIV/AIDS, who would have increased risk of stigmatization if their text message were viewed inadvertently, unless the message was encrypted. This was particularly of concern in South Africa, where health workers were repeatedly robbed at gunpoint for their mobile phones while making their home care visits [71].

Finally, as an unintended consequence of the high reach afforded and achieved using SMS, health workers may bear the larger burden of servicing a larger number of clients than before, creating problems in quality or in feelings of not being properly compensated [27]. This may be a challenge to programs that do not have the resources to enhance financial incentives of workers. Thus, careful planning about capacity of systems and their staff must be fully assessed prior to launching SMS initiatives that aim to increase use of services, addressing the potential increase in the staff workload.

Insights provided in this review outlined current practice of SMS for disease prevention in developing countries. Novel projects were doing innovative things to prevent the spread or progression of disease. Using games and interactive communication strategies, as well as integrating SMS with other channels, is a reality and holds great promise for reach and potential social change. Yet high-quality SMS-based intervention studies from developing countries were lacking in the literature, as reported in other recent literature reviews [72,73]. Most of the evaluations in this study were feasibility studies assessing process and output, rather than outcomes, a situation that has been called the pilot syndrome [74]. The included case studies predominantly appraised process and usage output and encompassed various limitations, such as (1) sample selection (no randomization; often convenience samples), (2) sample size (insufficient statistical power), (3) lack of information on process validations (pretesting, recruitment type, response rate, retention rate), (4) lack of assessment of the maintenance of the effect, and (5) no comparison with control groups. Given the enormous potential of this widespread communication tool, randomized controlled trials to determine efficacy of using SMS as a means for disease prevention in developing countries are highly needed and should be a priority in funding.

Limitations related to the literature review process should be considered when interpreting the results. The innovative aspects and the commercial implications of the field may have affected the reporting of studies. The findings could be influenced by a tendency of projects to report positive results, by time lag issues typical of a fast-moving field, by projects promoted by industry, or by unwillingness to share information in order to protect innovation. Additionally, this review includes SMS applications focused on disease prevention and thus does not provide a full picture of applications with other foci, such as disease surveillance or adherence to prescribed regimens. Results may furthermore have been influenced by the involuntary omission of documents from the gray literature, which were sometimes difficult to find [18]. The extensive search and the methodical approach used in this review strived to minimize limitations and enabled the assembly of comprehensive information about this innovative field.

Conclusion

There are many SMS health initiatives for disease prevention in developing countries, yet few are being evaluated and reported. Those that did conduct evaluations reported process evaluation and uptake, providing limited data about behavior change. Moreover, with a low number of documents found in the peer-reviewed literature, it appears that, to date, little is being done to advance our understanding of what works and what outcomes could be achieved in using SMS for disease prevention in a developing country contexts. Major opportunities are perceived, evident by the number and wide variety of projects, the recent creation of the United Nations Foundation's mHealth Alliance, and papers published describing SMS applications in the developed world [6-14,75]. However, the need remains for evidence-based dissemination of information about using mobile phones and SMS for improving health in the developing world. The limited evidence found in this systematic review highlights the need for research that assesses behavioral, social, economic, and health outcomes of mobile phone interventions aimed at promoting health in developing country contexts. Sharing best practices and providing strategic directions will allow the building of evidence for the use of mobile health technologies that address the needs of health systems, communities, and populations. Consequently, the promises of using mobile phones and SMS may translate into an equitable improvement in health.

Acknowledgments

The authors thank Mr Joel Lehmann for his assistance with protocol development and data collection, Mrs Sepideh Dianat for feedback on previous versions, and Mr Kaspar Hurni, Centre for Development and Environment, University of Berne, for establishing the world map figure.

Authors' Contributions

This research was undertaken as a collaborative project with the three authors.

Conflicts of Interest

None declared.

References

1. International Telecommunication Union. Measuring the Information Society: The ICT Development Index. Geneva: ITU; 2009.
2. GSM Association. GSMA. 2008. Universal Access: How Mobile can Bring Communications to All URL: http://www.gsmworld.com/documents/universal_access_full_report.pdf [WebCite Cache ID http://www.webcitation.org/63mos45N4]
3. Wikipedia. 2011 Mar 28. SMS URL: http://en.wikipedia.org/wiki/SMS [accessed 2011-03-29] [WebCite Cache ID 5xXmiDvlJ]
4. Kaplan WA. Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries? Global Health 2006;2:9 [FREE Full text] [doi: 10.1186/1744-8603-2-9] [Medline: 16719925]
5. Shields T, Chetley A, Davies J. infoDev. 2008. Improving Health, Connecting People: The Role of ICT in the Health Sector in Developing Countries URL: http://www.infodev.org/en/Document.486.html [accessed 2011-03-29] [WebCite Cache ID 5xXoRuX6x]
6. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
7. 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]
8. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
9. Vahatalo M, Virtamo H, Viikari J, Ronnemaa T. Cellular phone transferred self blood glucose monitoring: prerequisites for positive out-come. Practical Diabetes International 2004;21(5):192-194. [doi: 10.1002/pdi.642]
10. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
11. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
12. Kollmann A, Riedl M, Kastner P, Schreier G, Ludvik B. Feasibility of a mobile phone-based data service for functional insulin treatment of type 1 diabetes mellitus patients. J Med Internet Res 2007;9(5):e36 [FREE Full text] [doi: 10.2196/jmir.9.5.e36] [Medline: 18166525]
13. Joo NS, Kim BT. Mobile phone short message service messaging for behaviour modification in a community-based weight control programme in Korea. J Telemed Telecare 2007;13(8):416-420. [doi: 10.1258/135763307783064331] [Medline: 18078554]
14. Waugaman A. United Nations Foundation. 2010. mHealth Alliance URL: http://www.unfoundation.org/global-issues/technology/mhealth-alliance.html [accessed 2011-03-29] [WebCite Cache ID 5xXocEU7P]
15. Breslow L. From disease prevention to health promotion. JAMA 1999 Mar 17;281(11):1030-1033. [Medline: 10086439]
16. World Health Organization. WHO. 2011. The Ottawa Charter for Health Promotion URL: http://www.who.int/healthpromotion/conferences/previous/ottawa/en/index.html [accessed 2011-03-29] [WebCite Cache ID 5xXp34wbR]
17. World Health Organization. WHO. 2005 Jan 23. Global Health Promotion Scaling Up for 2015: A Brief Review of Major Impacts and Developments Over the Past 20 Years and Challenges For 2015 URL: http://www.who.int/healthpromotion/conferences/6gchp/hpr_conference_background.pdf [accessed 2011-03-29] [WebCite Cache ID 5xXpISYS4]
18. Dobbins P, Robeson PA. Health-evidence.ca. 2006 Mar 31. Methodology for Searching the Grey Literature for Effectiveness Evidence Syntheses Related to Public Health URL: http://health-evidence.ca/downloads/Dobbins2006_Grey_Literature_Project_Report.pdf [WebCite Cache ID http://www.webcitation.org/63mnwVHM0]
19. Howes F, Doyle J, Jackson N, Waters E. Evidence-based public health: The importance of finding 'difficult to locate' public health and health promotion intervention studies for systematic reviews. J Public Health (Oxf) 2004 Mar;26(1):101-104 [FREE Full text] [Medline: 15044584]
20. International Monetary Fund. IMF. 2008 Apr. World Economic and Financial Surveys: World Economic Outlook Database: WEO Groups and Aggregates Information: Country Composition of WEO Groups URL: http://www.imf.org/external/pubs/ft/weo/2008/01/weodata/groups.htm [accessed 2011-03-29] [WebCite Cache ID 5xXpYXEdG]

21. Higgins JPT, Green S, editors. The Cochrane Collaboration. 2008 Feb 29. Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 [updated February 2008] URL: http://www.mrc-bsu.cam.ac.uk/cochrane/handbook500/ [accessed 2011-12-06] [WebCite Cache ID 63jOCvuge]
22. African Broadcast Media Partnership Against HIV/AIDS. African Broadcast Media. 2009. It Begins With You URL: http://www.itbeginswithyou.org/ [accessed 2011-03-29] [WebCite Cache ID 5xXeg5yOM]
23. Freedom HIV/AIDS. ZMQ. 2008. Africa Reach Program URL: http://www.freedomhivaids.in/AfricaProgram.htm [accessed 2011-03-28] [WebCite Cache ID 5xWWuQOSQ]
24. United Nations Development Programme Botswana. UNDP. 2004. UNDP Botswana TV Programme Wins an Award. Sep URL: http://www.unbotswana.org.bw/undp/news_230904.html [accessed 2011-03-29] [WebCite Cache ID 5xXfCPHAL]
25. Townsend D. UNICEF West and Central Africa. 2009 Jan 6. Campaign Addresses Three Leading Causes of Child Mortality in the Central African Republic URL: http://www.unicef.org/wcaro/2009_2901.html [accessed 2011-12-09] [WebCite Cache ID 63nW9gMbI]
26. Shillingi LM, Mayo C. The Communication Initiative Network. 2009 Jul 9. Text Me! Flash Me! Helpline URL: http://www.comminit.com/en/node/291748/307 [accessed 2011-03-28] [WebCite Cache ID 5xWSYlImJ]
27. Clemmons L, Shillingi L. The Communication Initiative Network. 2009 Apr 1. Special Report: ICT is Effective in Supporting Behavior Change Communication, Strengthening Peer Outreach and Facility-Based Services, and Increasing Service Uptake among MSM and FSW URL: http://www.comminit.com/en/node/291706/2674 [accessed 2011-03-28] [WebCite Cache ID 5xWSnJ3tf]
28. Martin S. The Communication Initiative. 2011. eQuest URL: http://www.comminit.com/en/node/129640/2781 [accessed 2011-03-28] [WebCite Cache ID 5xWbxOfpM]
29. Makutano Junction. The Mediae Company. 2010. Welcome to the Makutano Junction Project URL: http://www.makutanojunction.org.uk/ [accessed 2011-03-28] [WebCite Cache ID 5xWVaEwqC]
30. Campbell D. The Communication Initiative Network. 2009. Makutano Junction URL: http://www.comminit.com/en/node/133770/306 [accessed 2011-03-28] [WebCite Cache ID 5xWVnIgG6]
31. Mobile for Good. OneWorld UK. 2011 Mar 28. M4G In Kenya: Services URL: http://uk.oneworld.net/article/view/117284 [accessed 2011-03-28] [WebCite Cache ID 5xWUWROkB]
32. African Network for Health Knowledge Management and Communication. Welcome to AfriAfya URL: http://www.afriafya.org/index.php [accessed 2011-03-29] [WebCite Cache ID 5xXfnABuh]
33. Netsquared. 2010 Dec 2. Pariah News: Where Citizen Media Enables AIDS Treatment Access URL: http://www.netsquared.org/projects/pariah-news-where-citizen-media-enables-aids-treatment-access [accessed 2011-03-29] [WebCite Cache ID 5xXg7G829]
34. PSI/Mali. The Communication Initiative Network. 2009 Apr 17. La Sante au Bout du Fil (Health On Line) URL: http://www.comminit.com/en/node/130505/307 [accessed 2011-03-28] [WebCite Cache ID 5xWRHVYPd]
35. Ofomata U. OneWorld UK. 2011 Mar 28. Learning About Living URL: http://uk.oneworld.net/article/archive/9789 [accessed 2011-03-28] [WebCite Cache ID 5xWbnORGm]
36. Ofomata U, Groves A. Oneworld.net. 2009 Apr. Learning About Living: Lessons Learnt: 2007-2009 URL: http://mobile.oneworld.net/docs/lal/Lessons_Learnt_Nigeria_2009.pdf [accessed 2011-03-28] [WebCite Cache ID 5xWcMCR36]
37. Stepanek M. Mobileactive.org. 2009 Mar 15. 63 Million Bednets to be Distributed with RapidSMS URL: http://mobileactive.org/net-speed-unicef-distributes-70-million-bednets-nigeria-sms [accessed 2011-03-28] [WebCite Cache ID 5xWNXjLy0]
38. Community Media Trust. 2010. Siyayinqoba Beat It! URL: http://www.beatit.co.za/home/index [accessed 2011-03-28] [WebCite Cache ID 5xWb76gde]
39. Communication Health Media Trust, Phooko P. The Communication Initiative Network. 2009 Apr 20. Siyayinqoba (Beat It) URL: http://www.comminit.com/en/node/128365/347 [accessed 2011-03-28] [WebCite Cache ID 5xWbR64EL]
40. Cell-life.org. 2009. Cell Phones 4 HIV URL: http://www.cell-life.org/cellphones-4-hiv/ [accessed 2011-03-28] [WebCite Cache ID 5xWO3m6L3]
41. Skinner D. Cell-life.org. Evaluation of Cell Phone for Life for Cell-Life: Baseline Report URL: http://www.cell-life.org/images/stories/file%202%20-%20donald%20skinner%20baseline%20report%20tac%20wc.pdf [accessed 2011-03-28] [WebCite Cache ID 5xWJ2lIbp]
42. Butcher D. Mobile Marketer. 2008 Oct 30. Project M Delivers HIV/AIDS, TB Care Via Mobile URL: http://www.mobilemarketer.com/cms/news/messaging/2006.html [accessed 2011-03-28] [WebCite Cache ID 5xWOJLY9P]
43. Eichenseher T. National Geographic. 2008. Text Messages Used to Combat AIDS in S. Africa URL: http://news.nationalgeographic.com/news/2008/10/081024-cellphone-HIV.html [accessed 2011-03-28] [WebCite Cache ID 5xWPpTGT4]
44. SADAG. South African Depression and Anxiety Group URL: http://www.sadag.co.za/ [accessed 2011-03-28] [WebCite Cache ID 5xWWYwdfD]
45. Veldhuizen J, Harmsen S. IICD. 2009 Apr 15. Learning About the Many Ways of Using Mobile Phones in Health URL: http://www.iicd.org/articles/learning-about-the-many-ways-of-using-mobile-phones-in-health/?searchterm=SMS [accessed 2011-03-29] [WebCite Cache ID 5xXhSUjGV]
46. Kimasomaso, BBC World Service Trust. The Communication Initiative Network. 2010 Jul 7. Kimasomaso URL: http://www.comminit.com/en/node/134365/303; [accessed 2011-03-29] [WebCite Cache ID 5xXj57P3N]

47. BBC World Service Trust. 2003. East Africa: Sexual Health Radio Show. Sep 9 URL: http://www.bbc.co.uk/worldservice/trust/projectsindepth/story/2003/09/030909_kimasomaso.shtml [accessed 2011-12-09] [WebCite Cache ID 63nUpadW0]
48. AppLab. Grameen Foundation. 2010. Health Applications Deployed: Clinic Finder and Health Tips (Powered by Google SMS) URL: http://www.grameenfoundation.applab.org/section/health-apps [accessed 2011-03-28] [WebCite Cache ID 5xWVM0Bzh]
49. Text to Change. Text To Change: Spread the Message, Stop the Virus URL: http://www.texttochange.org/ [accessed 2011-03-28] [WebCite Cache ID 5xWarqpFQ]
50. netsquared.org. 2000 Mar 9. Text to Change: HIV/AIDS Education, Data Gathering and Uptake of VCT Via Mobile Telephony URL: http://www.netsquared.org/projects/text-change-hivaids-education-data-gathering-and-uptake-vct-mobile-telephony-0 [accessed 2011-12-05] [WebCite Cache ID 63hZcCgcv]
51. Mitchell KJ, Bull S, Kiwanuka J, Ybarra ML. Cell phone usage among adolescents in Uganda: acceptability for relaying health information. Health Educ Res 2011 Oct;26(5):770-781. [doi: 10.1093/her/cyr022] [Medline: 21536715]
52. UNICEF. 2009 Feb 4. UNICEF Humanitarian Action Update: Zimbabwe URL: http://www.unicef.org/infobycountry/files/ZIMBABWE_HAU_4_Feb_2009.pdf [accessed 2011-12-09] [WebCite Cache ID 63nUwNLZk]
53. Textually.org. 2004 Sep 21. China Netcom Launches SMS Sexual Education Project URL: http://www.textually.org/textually/archives/2004/09/005412.htm [accessed 2011-03-29] [WebCite Cache ID 5xXkQDXsK]
54. Textually.org. 2004 Jan 10. SARS Text Alerts Reactivated URL: http://www.textually.org/textually/archives/2004/01/002758.htm [accessed 2011-03-29] [WebCite Cache ID 5xXkh0FPv]
55. Jardin X. Wired.com. 2003 Apr 21. Text Messaging Feeds SARS Rumors URL: http://www.wired.com/medtech/health/news/2003/04/58506 [accessed 2011-03-29] [WebCite Cache ID 5xXkwBxgE]
56. USAID Indonesia. United States Agency for International Development, Community Based Avian Influenza Control Project. 2008 Jan. Bird flu contest for travelers. The Avian Influenza Roundup, Issue no. 1, page 3 URL: http://www.dai.com/pdf/The_Avian_Influenza_Roundup,_No._1_-_ENG.pdf [accessed 2011-11-30] [WebCite Cache ID 63Znc9ReC]
57. BBC World Service Trust. 2008 Mar 7. Normalising Condoms in India URL: http://www.bbc.co.uk/worldservice/trust/asiapacific/story/2008/03/080307_jbws_phase2.shtml [accessed 2011-12-09] [WebCite Cache ID 63nUt0gpy]
58. MacPherson Y. BBC Internet Blog. 2008 Sep 9. World Service Trust Condom Ringtone URL: http://www.bbc.co.uk/blogs/bbcinternet/2008/09/world_service_condom_ringtone.html [accessed 2011-03-28] [WebCite Cache ID 5xWYWn8Mq]
59. ZMQ. 2009. Freedom HIV/AIDS URL: http://www.freedomhivaids.in/FreedomHivAids.htm [accessed 2011-03-28] [WebCite Cache ID 5xWWmIDKF]
60. Heroes AIDS Project. 2010. Heroes Project URL: http://www.heroesprojectindia.org/index.asp [accessed 2011-03-31] [WebCite Cache ID 5xbF82iWS]
61. Andhra Cafe. 2007 Mar 17. Query on AIDS? SMS 7827 for Help URL: http://andhracafe.com/index.php?m=show&id=20306 [WebCite Cache ID http://www.webcitation.org/63mobMLQg]
62. Textually.org. 2004 Mar 22. TB Public Awareness Campaigns in India URL: http://www.textually.org/textually/archives/2004/03/003372.htm [accessed 2011-03-29] [WebCite Cache ID 5xXjlW7kR]
63. Khokhar A. Short text messages (SMS) as a reminder system for making working women from Delhi Breast Aware. Asian Pac J Cancer Prev 2009;10(2):319-322 [FREE Full text] [Medline: 19537904]
64. BBC News. 2008 Oct 15. Millions Mark UN Hand-Washing Day URL: http://news.bbc.co.uk/2/hi/south_asia/7670855.stm [accessed 2011-03-28] [WebCite Cache ID 5xWLc6xZz]
65. Maqbool S. The News International. 2009 Apr 11. Mobile Phone Company Starts Sending Anti-polio Text Messages URL: http://www.thenews.com.pk/TodaysPrintDetail.aspx?ID=171938&Cat=6&dt=4/11/2009 [WebCite Cache ID http://www.webcitation.org/63mp4LDFs]
66. Sex-Ed Text: Text Hotline For Reproductive Health & Sex Education. What is SET? URL: http://www.freewebs.com/sexedtext/ [accessed 2011-03-28] [WebCite Cache ID 5xWV6DB7q]
67. Instituto Carlos Slim de la Salud. Voxiva. CardioNet URL: http://www.voxiva.com/cardionet/index.html [accessed 2011-03-28] [WebCite Cache ID 5xWaErVID]
68. OneWorld UK. Learning About living URL: http://www.learningaboutliving.org/south [accessed 2011-09-29] [WebCite Cache ID 623hlpMQp]
69. Brand South Africa. The Languages of South Africa URL: http://www.southafrica.info/about/people/language.htm [accessed 2011-03-29] [WebCite Cache ID 5xXqaqVgx]
70. Crespo Garcia D. Information Technology and Communication Technologies for Development (ICT4D). 2008. Mobile Phones' Potential to Address Information and Communication Needs of Healthcare Workers in Isolated Rural Areas in Peru URL: http://www.gg.rhul.ac.uk/ict4d/workingpapers/crespo.pdf [accessed 2011-03-29] [WebCite Cache ID 5xXqsBVPT]
71. Lindow M. Wired.com. 2004 Apr 11. How SMS Could Save Your Life URL: http://www.wired.com/medtech/health/news/2004/11/65585 [accessed 2011-03-29] [WebCite Cache ID 5xXr23zVB]
72. Noordam AC, Kuepper BM, Stekelenburg J, Milen A. Improvement of maternal health services through the use of mobile phones. Trop Med Int Health 2011 May;16(5):622-626. [doi: 10.1111/j.1365-3156.2011.02747.x] [Medline: 21342374]

73. Mukund Bahadur KC, Murray PJ. Cell phone short messaging service (SMS) for HIV/AIDS in South Africa: a literature review. Stud Health Technol Inform 2010;160(Pt 1):530-534. [Medline: 20841743]
74. Chetley A, editor. infoDev. 2006 May 31. Improving Health, Connecting People: The Role of ICTs in the Health Sector of Developing Countries: A Framework Paper (Working paper no. 7) URL: http://www.infodev.org/en/Document.84.html [accessed 2011-03-29] [WebCite Cache ID 5xXrcvzOT]
75. Fox S. Mobile health 2010. Washington, DC: Pew Internet & American Life Project; 2010 Oct 19. URL: http://www.pewinternet.org/~/media//Files/Reports/2010/PIP_Mobile_Health_2010.pdf [accessed 2011-03-31] [WebCite Cache ID 5xbO4J2J6]

Abbreviations

HIV/AIDS: human immunodeficiency virus/acquired immunodeficiency syndrome
ICT: information and communication technologies
SMS: short message service
SRH: sexual and reproductive health
VCT: voluntary counseling and testing

Edited by G Eysenbach; submitted 01.04.11; peer-reviewed by R Furberg, H Thinyane, J Barrington, G Wright, E Kuntsche; comments to author 23.04.11; revised version received 03.10.11; accepted 05.10.11; published 12.01.12.

Please cite as:
Déglise C, Suggs LS, Odermatt P
Short Message Service (SMS) Applications for Disease Prevention in Developing Countries
J Med Internet Res 2012;14(1):e3
URL: http://www.jmir.org/2012/1/e3/
doi:10.2196/jmir.1823
PMID:22262730

Original Paper

MEMO—A Mobile Phone Depression Prevention Intervention for Adolescents: Development Process and Postprogram Findings on Acceptability From a Randomized Controlled Trial

Robyn Whittaker[1], MBChB, MPH, PhD; Sally Merry[2], MBChB, MD; Karolina Stasiak[2], PhD; Heather McDowell[3], PhD; Iain Doherty[4], PhD; Matthew Shepherd[2], BA(Hons); Enid Dorey[1], MSc; Varsha Parag[1], MSc; Shanthi Ameratunga[5], MBChB, MPH, PhD; Anthony Rodgers[6], MBChB, PhD

[1]Clinical Trials Research Unit, University of Auckland, Auckland, New Zealand

[2]Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

[3]Auckland District Health Board, Auckland, New Zealand

[4]Learning Technology Unit, University of Auckland, Auckland, New Zealand

[5]Injury Prevention Research Unit, University of Auckland, Auckland, New Zealand

[6]George Institute for Global Health, University of Sydney, Sydney, Australia

Corresponding Author:
Robyn Whittaker, MBChB, MPH, PhD
Clinical Trials Research Unit
University of Auckland
School of Population Health
Tamaki Campus
Auckland, 1042
New Zealand
Phone: 64 93737599
Fax: 64 93731710
Email: r.whittaker@ctru.auckland.ac.nz

Abstract

Background: Prevention of the onset of depression in adolescence may prevent social dysfunction, teenage pregnancy, substance abuse, suicide, and mental health conditions in adulthood. New technologies allow delivery of prevention programs scalable to large and disparate populations.

Objective: To develop and test the novel mobile phone delivery of a depression prevention intervention for adolescents. We describe the development of the intervention and the results of participants' self-reported satisfaction with the intervention.

Methods: The intervention was developed from 15 key messages derived from cognitive behavioral therapy (CBT). The program was fully automated and delivered in 2 mobile phone messages/day for 9 weeks, with a mixture of text, video, and cartoon messages and a mobile website. Delivery modalities were guided by social cognitive theory and marketing principles. The intervention was compared with an attention control program of the same number and types of messages on different topics. A double-blind randomized controlled trial was undertaken in high schools in Auckland, New Zealand, from June 2009 to April 2011.

Results: A total of 1348 students (13–17 years of age) volunteered to participate at group sessions in schools, and 855 were eventually randomly assigned to groups. Of these, 835 (97.7%) self-completed follow-up questionnaires at postprogram interviews on satisfaction, perceived usefulness, and adherence to the intervention. Over three-quarters of participants viewed at least half of the messages and 90.7% (379/418) in the intervention group reported they would refer the program to a friend. Intervention group participants said the intervention helped them to be more positive (279/418, 66.7%) and to get rid of negative thoughts (210/418, 50.2%)—significantly higher than proportions in the control group.

Conclusions: Key messages from CBT can be delivered by mobile phone, and young people report that these are helpful. Change in clinician-rated depression symptom scores from baseline to 12 months, yet to be completed, will provide evidence on the effectiveness of the intervention. If proven effective, this form of delivery may be useful in many countries lacking widespread mental health services but with extensive mobile phone coverage.

ClinicalTrial: Australia New Zealand Clinical Trials Registry (ACTRN): 12609000405213; http://www.anzctr.org.au/trial_view.aspx?ID=83667 (Archived by WebCite at http://www.webcitation.org/64aueRqOb)

(J Med Internet Res 2012;14(1):e13) doi:10.2196/jmir.1857

KEYWORDS

Mobile phone; depression prevention; adolescents; cognitive behavioral therapy

Introduction

In 1990, depressive disorder was ranked fourth in the estimate of global disease burden. It has been predicted that it will be the second most important cause of lost healthy life-years globally by 2020 [1]. Depressive disorder commonly starts in adolescence, and its effect on young people is pervasive with respect to overall development [2-5]. It is associated with poor academic functioning, social dysfunction, substance use, and attempted and completed suicide [2,6,7]. Comorbidity is high, with up to half of those with major depressive disorder having a lifetime occurrence of another psychiatric disorder. In New Zealand, depressive disorder is a major health issue among adolescents with rates of 4%–8% at the age of 15 years rising rapidly to 17%–18% by the age of 18 years [8]. Rates are higher for young people of Maori (the indigenous population of New Zealand), Pacific, and some Asian ethnic groups [9-11].

The marked increase in period prevalence estimates from mid to late adolescence makes this a good time to intervene to prevent the onset of depressive disorder. Psychological interventions such as cognitive behavioral therapy (CBT) have been shown to be an effective treatment for depression in adolescents [12,13] and may be effective in preventing the onset of depressive disorder in children and adolescents [14]. In this review of predominantly CBT-based interventions, the change in depression scores was modest; however, the number needed to treat (NNT) to prevent one adolescent developing a depressive disorder was approximately 10. In comparison, over 800 people with uncomplicated hypertension must be treated for a year to prevent one stroke, and 67 people who have had a myocardial infarction need to take aspirin to prevent one death from any cause [15,16]. Reduction in depression has been shown for both interventions targeted at those at risk (NNT of 8) and universal interventions for all adolescents (NNT of 13). The most effective study reported to date was a program targeted at children whose parents had mood disorders, where the NNT was 4 [17].

A randomized placebo-controlled trial of a school-based depression prevention program (RAP-Kiwi) used manuals with instructions on CBT in graphics and text, and comprised 11 sessions delivered during the school day as one of the regular health classes for all students [18]. A total of 392 participants were recruited from two different years in two schools in Auckland, New Zealand—one from a lower socioeconomic urban area and the other from a middle-class rural district. Immediately after the program, depression scores (measured by the Beck Depression Inventory and Reynolds Adolescent Depression Scale, Second Edition [RADS-2]) were reduced significantly more by RAP-Kiwi than by placebo. Categorical analysis confirmed significant clinical benefit with an absolute risk reduction of 3% (95% confidence interval 1%–11%, McNemar χ^2, P = .03) and an NNT for short-term benefit of 33. However, the considerable teaching time required and the difficulty ensuring fidelity of the program were both considered problematic with respect to a widespread rollout. These factors suggested that a new delivery medium would be necessary.

The ubiquitous spread of mobile phones allowed the consideration of an automated intervention for widespread cost-effective dissemination. Mobile phones have been successfully used to deliver behavior change interventions for smoking cessation [19], medication reminders [20-22], diet and physical activity [23-26], and the management of diabetes [27-30]. Our research team has used behavior change theory and techniques to guide the development of mobile phone interventions using text and video messages [31-33]. This research brought together this mobile health expertise with expertise from the RAP-Kiwi depression prevention research. We set out to adapt traditionally manualized and therapist-delivered CBT techniques into automated messages that could be delivered on a multimedia-capable mobile phone, with the aim of preventing the onset of depressive disorder in all New Zealand adolescents.

Objective

The objective of the randomized controlled trial (RCT) was to test whether the mobile phone CBT intervention can improve subjective and objective scores of depression symptoms in adolescents in comparison with a control group at 12 months (to be published separately). The aim of this paper is to describe the developmental process, and the acceptability and utility of the intervention as reported by the adolescents at the end of the intervention (9 weeks).

Methods

Intervention Development

Initial focus groups with students (n = 27) in a low socioeconomic multicultural high school formed an understanding of how adolescents use their mobile phones and how a well-being and problem-solving program could be useful and appealing to them. All participants used mobile phones, predominantly for text messaging. Concerns were around the cost of other functions beyond text messaging and potential loss of confidentiality with video calling. In general, students felt that text messages would be useful for information and positive reinforcement, videos could be used to demonstrate strategies for dealing with problems, and music or music videos could be used for relaxation. Some initial short videos were pretested with students (n = 40) at a small predominantly indigenous high school via computer. The students' preferences for different styles of videos and animations directly informed the

development of the intervention. Once some initial content had been developed, pretesting took place during the development process with groups of adolescents (n = 28) in various environments (youth centers, libraries, and invited group sessions). A variety of different text and video messages were shown to participants on a laptop computer. Their feedback supported the use of a variety of message types, as different young people found different mediums of delivery appealing. Important themes arising from these discussions were the importance of realism, credibility, positivity, and simple, clear messages.

An expert content group with expertise in adolescent psychiatry and psychological therapies, CBT, learning technology, marketing and media, and mobile phone health interventions was formed. The group derived 15 key messages, considered appropriate for delivery by mobile phone, from CBT and from RAP-Kiwi (Textbox 1).

To enhance relevance and thereby engagement with the program, it was important to deliver these key CBT themes within everyday contexts and addressing common issues faced by New Zealand adolescents. This was also a good fit with the theoretical basis, derived from social cognitive theory and tested in our previous smoking cessation video messaging intervention [33], of enhancing self-efficacy to deal with issues and life events using cognitive techniques [34]. This led to the use of observational learning to observe other "ordinary" teens facing typical issues or events and using the techniques to feel more positive about these issues. The idea was that the observers are more likely to try the techniques if they see other people like themselves successfully using them [35-40]. For example, story lines delivered in a series of short videos illustrate how moods follow thoughts, with negative thoughts leading to low mood, and how the recognition of cognitive distortions (often by a friend) can lead to a more positive mood. Actual techniques, such as a problem-solving mnemonic (STEPS), could be normalized in the demonstration of its use by adolescents for common adolescent problems.

The mode of delivery was carefully considered: whereas some Web- and computer-based interventions have been primarily manuals adapted for use on the computer, this approach has been associated with poor adherence and was unlikely to work for mobile phones [41]. A variety of types of mobile phone messages were considered—video clips, text messages, and animated cartoons—in order to widen the appeal and potential for engagement with the target audience. In this way, the same message could be repeated in different formats to increase recall.

As well as the ordinary teen role models described above, our initial focus groups had confirmed the importance of celebrities for teenagers, and so we included video messages from celebrities. In these short video clips they discussed the same strategies for dealing with difficult times in their own lives. Animations were also popular with our formative research participants. A cartoon was developed with *mobisodes* (short episodes by mobile phone) about four fictitious adolescents—repeating the key CBT messages in a different format within a different story that might appeal to a different cohort of teenagers. Homework, trying out the techniques in their own lives, is an important part of CBT and was addressed by celebrity video messages setting relevant weekly challenges for participants.

Based on the feedback from adolescents in the focus groups and pretesting, a regimen of 2 messages per day (outside school hours) over 9 weeks was developed. Media and marketing principles were applied to the importance of overall framing of the program and to the careful design of individual text and video messages. The program was given identity and coherence with a logo and byline (MEMO: living in a positive space), theme music, and three key words (spot, sort, do). The key CBT messages were reinforced by short memorable text messages that all related to one of the three key words.

To provide ongoing contact and reminders of the themes after the 9-week intervention, monthly text messages directed participants to a mobile website. This website (available only to recognized participant mobile phone numbers) provided a summary of the key messages, information on how to get more help, and a downloadable relaxation audio. New videos were posted on the website monthly. Ringtones, wallpaper images, and music downloads were linked to the mobile website.

Textbox 1. The 15 key cognitive behavioral therapy-based messages that formed the basis of the MEMO intervention

Your feelings are a result of what you think and what you do
You can take control of this
Being busy increases happiness generally
Do fun stuff
Don't procrastinate
Relaxation makes you feel good
Being with people you like makes you feel good
Fighting with people makes you feel bad
Sometimes you need to ask for help
Noticing only weaknesses and failures makes you feel bad
It's not what happens; it's what you think about it that affects feelings
We can choose to look at the world in a positive or a negative way
We can deal with negative thoughts
There are ways of dealing with stress
There are problem-solving techniques

Study Design

This research was part of a double-blind RCT (June 2009–April 2011) designed to test the effectiveness of the mobile phone intervention. Human subjects ethical approval was provided by the Northern Region Y Ministry of Health Human Ethics Committee (NTY/0/09/088).

Support from a mobile telecommunications company (Vodafone New Zealand Ltd, Auckland, New Zealand) ensured that the program was delivered completely free of cost to participants.

Recruitment

A total of 15 high schools across Auckland city (population approximately 1.4 million) agreed to participate in the study. The included schools represented a range of private/public, single-sex/coeducation, ethnicity, and decile ratings (a school's decile reflects the socioeconomic status of the school's community based on the New Zealand census). Study processes were agreed on with each school guidance counselor or team. Parents of potential participants were sent participant information sheets and study contact details, and were given the opportunity to opt their child out of the study. One school required written parental consent. The research team promoted the study to large groups of students at assemblies or other organized school sessions. It was described as a study involving a mobile phone program about "living in a positive space." Those who wished to participate were given full study information and were asked to complete a written consent form and baseline data collection forms immediately after the presentation. Students identified as having current depressive symptoms (score of 76 or higher on the self-completed RADS-2) or at risk of self-harm (question 14 on the RADS-2, "I feel like harming myself sometimes" or "often") were immediately referred for management by the school guidance counselors according to school protocols [42]. These students were excluded from this study, as the research question was about the prevention of onset of depressive disorder.

All students completing baseline data collection received a run-in program consisting of daily mobile phone messages for 9 days. These messages provided instructions about the study and allowed students to determine whether they could view the video messages on their mobile phones (if not, they could not proceed to randomization). This run-in period was also designed to allow for the initial sharing of messages between friends without potentially contaminating randomized groups. During this period, potential participants were invited to an individual interview with a trained research assistant who conducted the Child Depression Rating Scale-Revised (CDRS-R) [43]. Again, any students exhibiting current depression or risk of self-harm were referred for management by the school guidance counselors and excluded from the study.

At this point, all those who met the eligibility criteria were randomly assigned to receive either the intervention or control program. Allocation concealment was maintained by computer-based randomization so that researchers were unaware of possible allocation. A stratified minimization was used to ensure balance for possible prognostic factors: sex, ethnicity (Maori/Pacific vs non-Maori/non-Pacific), and school. Active recruitment ceased once the target sample size had been reached—all students who had been offered participation at that stage were accepted into the study.

Intervention and Control

Students in the intervention group then received 2 messages per day for 9 weeks (outside school hours), followed by monthly messages and access to a mobile website. The messages were a mixture of text messages, video messages of adolescents and celebrities, and animated cartoons (described above; see Multimedia Appendix 1 for an example video message). Because of the potential for high levels of placebo response in depression interventions, students in the control group received a full

attention control program with the same number of mobile phone messages, the same types of messages, including the same adolescents and celebrities in the video messages, and the same characters in the cartoon messages. The content of the placebo messages was focused on healthy eating, sustainability of the environment, and safe practices for using the Internet and mobile phone (cybersafety). Participants were not aware of which program was the intervention and which was the control.

Outcome Measures

Individual interviews were arranged with participants at the end of the 9-week program and took place on school grounds. During these interviews, participants self-completed forms on how much of the program they had viewed (as a measure of adherence to the program) and their perceived usefulness of the program. The interviews were conducted by research assistants blinded to allocation. The primary outcome for the trial will be the change in the clinician-assessed (blinded) depression symptom scores (CDRS-R) from baseline to 12 months [43]. Secondary outcomes will include change in the self-rated RADS-2 depression symptom scale [42]; incidence of depression; general and school functioning; and quality of life.

Statistical Analysis

The target sample size for the trial was calculated for the primary RCT outcome of change in CDRS-R (not reported here) at 790 participants. This was based on randomization at the individual level, as cluster randomization by school would have required a larger sample that was not considered feasible by the research team. All analyses were 2-tailed and conducted according to a prespecified plan using SAS version 9.2 (SAS Institute, Cary, NC, USA). Continuous variables are presented as means and standard deviations, and categorical variables are presented as counts and percentages. Differences between groups for categorical outcomes were analyzed using chi-square tests or Fisher exact test.

Results

Figure 1 depicts the flow of participants through study processes. From an estimated 2850 students attending the presentation sessions, 1348 students (47.30%) registered initial interest in participating in the study. In three cases, parents actively opted their child out of the study. Of those registered, 122 (9.1%) were referred to the school guidance counselors for immediate management of depression or risk of self-harm. A further 371 (27.5%) dropped out at this stage due to lack of interest or difficulty viewing the video messages on their phones, or were excluded for having been recently through a school-based depression program for at-risk students. A total of 855 students (63.4% of those registered) were individually randomly assigned: 426 to the intervention group and 429 to the control group.

Of the 855 randomly assigned participants, 584 (68.3%) were female and 271 (31.7%) were male (Table 1). The largest group was New Zealand Europeans (501, 58.6%), followed by Asian students (208, 24.3%), Maori (83, 10%; the indigenous population of New Zealand who make up approximately 14% of the overall population), and Pacific students (51, 6%). This gender and ethnicity distribution mostly reflects the participating schools—for example, overall the schools in the study contained 9% Maori students and 5% Pacific students. The age range of participants is from 13 years to 17 years, with a mean age of 14 years.

Participants were asked to specify (within categories) how many of the messages they viewed. Approximately three-quarters of the intervention group (311/418, 74.4%) viewed at least half of the messages, with 29.6% (n = 123) viewing most or all of the messages (Table 2). Although more than a third (324/835, 38.8%) of participants shared messages (with anyone, not necessarily someone in the study), they mostly shared only a small number (<10) of messages. This number of messages was not considered sufficient to have an effect even if shared with someone in the other group (contamination). Participants in the intervention group seemed to like the types of messages that were used: 78.0% (326/418) liked the video messages from celebrities; 71.8% (300/418) liked the video messages from other teens; 65.8% (275/418) liked the animated cartoons; and 60.3% (252/418) liked the text messages.

Figure 1. CONSORT flowchart of participants in the MEMO trial.

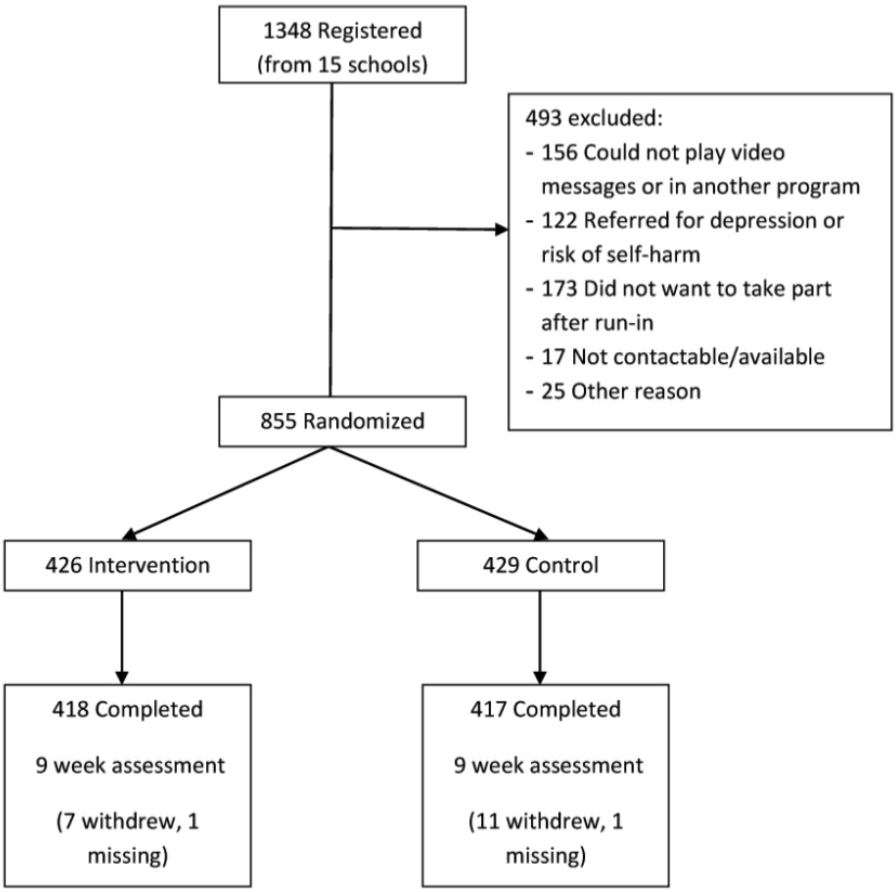


Table 1. Baseline participant characteristics

	Intervention (n = 426)	Control (n = 429)
Sex, n (%)		
Female	291 (68.3)	293 (68.3)
Male	135 (31.7)	136 (31.7)
Ethnicity, n (%)		
NZ[a] European	245 (57.5)	256 (59.7)
Maori (NZ indigenous population)	38 (9)	45 (11)
Asian	113 (26.5)	95 (22)
Pacific	23 (5)	28 (7)
Other	7 (2)	5 (1)
Age (years), mean (SD)	14.3 (0.90)	14.3 (0.91)
School year, n (%)		
Year 9 (13 & 14 years of age)	74 (17)	84 (20)
Year 10 (14 & 15 years of age)	174 (40.9)	169 (39.4)
Year 11 (15 & 16 years of age)	148 (34.7)	148 (34.5)
Year 12 (16 & 17 years of age)	30 (7)	28 (7)

[a] New Zealand.

Table 2. Participant viewing and sharing of messages

	Intervention (n = 418)	Control (n = 417)
Number of messages viewed, n (%)		
Most/all	123 (29.6)	142 (34.4)
More than half	96 (23)	110 (26.6)
About half	92 (22)	82 (20)
Some	74 (18)	56 (14)
Hardly any	30 (7)	24 (6)
Missing data	3	3
Number of messages shared with others, n (%)	158 (37.9)	166 (40.1)
1–9	111 (70.3)	125 (75.8)
10–19	37 (23)	31 (19)
20–29	5 (3)	7 (4)
30+	5 (3)	2 (1)

The majority of participants (688/835, 82.4%) said they found MEMO to be useful (Table 3). A significantly greater proportion of those in the intervention group (379/418, 90.7%) said they would recommend it to a friend than those in the control group (345/417, 82.7%, $P < .001$). Participants were asked whether the program helped them with particular topics—some topics from the intervention and some from the control program. More participants in the intervention group than in the control group said that MEMO helped them to be more positive, to get rid of negative thoughts, to relax, to solve problems, to have fun, and to deal with issues in school (Table 3). These make up some of the key CBT messages in the intervention. As expected, more control group participants said the program helped them with topics covered in the control program (eg, be safe on the Internet, think more about the environment). Similar proportions in both groups said the program helped them with some of the issues considered to be in the intervention only, such as ‘to help other people,” and disappointingly the intervention group participants were no more likely than the control group participants to know where to go for help.

Table 3. Participant satisfaction and perceived usefulness of MEMO

	Intervention (n = 418)	Control (n = 417)	χ^2 *P* value
Found MEMO to be helpful, n (%)	351 (84.0)	337 (80.8)	.29
Would recommend MEMO to friends, n (%)	379 (90.7)	345 (82.7)	<.001
MEMO helped me to..., n (%)			
be more positive	279 (66.7)	209 (50.1)	<.001
be nicer to people	140 (33.5)	102 (24.5)	.004
get rid of negative thoughts	210 (50.2)	135 (32.4)	<.001
relax	222 (53.1)	176 (42.2)	.002
solve problems	138 (33.0)	104 (24.9)	.01
have fun	186 (44.5)	160 (38.4)	.07
deal with issues at school	109 (26.1)	80 (19)	.02
deal with issues at home	88 (21)	68 (16)	.08
be healthy	112 (26.8)	233 (55.9)	<.001
be safe on the net	68 (16)	175 (42.0)	<.001
get support when I need it	122 (29.2)	94 (23)	.03
be safe with my mobile phone	101 (24.2)	179 (42.9)	<.001
know who to go to when I need help	137 (32.8)	139 (33.3)	.86
think more about the environment	75 (18)	176 (42.2)	<.001
help other people	177 (42.3)	160 (38.4)	.24
speak out about things I'm passionate about	145 (34.7)	104 (24.9)	.002

More female students than male students perceived the program as being helpful (Table 4). This appears to hold true for most of the key messages except where more male students stated "MEMO helped me to relax." Participants in the intervention group who answered that they did not find MEMO to be useful stated their reasons in free text. These were categorized as follows: it didn't change the way I thought (n = 15); I didn't have any problems/already happy so didn't need any help (n = 15); I had technical difficulties (n = 15, including not viewing the videos, taking too long to download, and lost or broken phone); issues covered were too minor or not relevant (n = 10); it was boring/nothing new (n = 7); other (n = 7). The intervention group's suggestions for improvements were overwhelmingly related to reducing the number of messages. Other suggestions included providing solutions to other (more serious) problems.

The intervention ran as intended throughout the study, but technical issues arose from the large number of overseas mobile phones (not distributed by New Zealand telecommunications companies) that did not have the appropriate New Zealand Internet settings. Research staff had to assist these students to change the settings before they could commence the program. Also the telecommunications company changed the way they charged customers for Internet access during the study period, which caused a few participants to be charged for viewing the video messages. This was addressed once the issue was identified, and the participants were reimbursed for any charges.

Table 4. Perceived usefulness by sex

	Intervention		Control	
	Female (n=285)	Male (n=133)	Female (n=287)	Male (n=130)
Would recommend MEMO to friends, n (%)	259 (91.0)	120 (90.2)	240 (83.6)	105 (80.8)
MEMO helped me to..., n (%)				
be more positive	205 (71.9)	74 (56)	149 (51.9)	59 (45)
get rid of negative thoughts	154 (54.0)	56 (42)	98 (34)	37 (29)
relax	142 (49.8)	80 (60)	124 (43.2)	52 (40)
solve problems	100 (35.1)	38 (29)	81 (28)	23 (18)
know who to go to when I need help	91 (32)	46 (35)	105 (36.6)	34 (26)
help other people	130 (45.6)	47 (35)	114 (39.7)	46 (35)

Discussion

This study shows that key messages from CBT can be delivered by mobile phone and that young people report that these are helpful. This is the first study to deliver a CBT-based intervention to prevent depression in adolescents via mobile phone text and video messages. Feedback from New Zealand adolescents has shown that the program is acceptable, with three-quarters watching more than half of the messages and a very high proportion recommending it to their friends. Even more encouraging are the findings that participants felt MEMO helped them to be more positive, get rid of negative thoughts, relax, solve problems, and deal with issues at school. More work could be done to refine the intervention for the future; in particular, participants suggested that the number of messages be reduced. Further qualitative research with adolescents is also underway to obtain more in-depth feedback on the intervention, and to determine its appeal to different subgroups and how to improve it.

Limitations of this study include the use of self-reported outcomes from adolescents. As participants did not know whether the program they received was the intervention or the control program, this is unlikely to have a major effect on the findings. We did not measure possible mediators such as self-efficacy (from the theoretical base), nor did we ask participants whether they had tried any of the techniques suggested. Qualitative research conducted after the RCT may address some of these and will be reported separately. It is not clear how generalizable these results will be for other populations. In particular, it is disappointing for us that the trial did not recruit a large proportion of young Maori participants. However, this was representative of the schools involved in the study, and in terms of absolute numbers this compares favorably with other studies of CBT-related interventions with indigenous populations. Two other major ethnic minority groups appear to be reasonably well represented (Pacific people comprise 6.9%, and Asian people 9.2%, of the New Zealand population, according to Statistics New Zealand).

Primary outcomes from the RCT will be required to determine the overall effectiveness of the intervention with respect to the prevention of depression symptoms at 12 months postrandomization. If the encouraging results presented here are maintained, mobile phone programs could be a cost-effective method for delivering basic CBT techniques to all adolescents. Such programs can be easily scaled up to reach large disparate populations regardless of geographic location, as has been shown by the implementation of a successful text message smoking cessation intervention [31] as a free national program in New Zealand [44]. Promotion through schools is one option, as shown in the study, but other distribution options may also be possible. Also, the development concepts and key messages in this intervention may translate to other populations with the adaptation of local content. With the increasing prevalence of mobile phone use in low- and middle-income countries and the use of mobile health technologies to address many aspects of noncommunicable diseases, using strategies of this kind has significant potential to address global disparities in the burden of adolescent depression.

Acknowledgments

This research was supported by the Health Research Council of New Zealand, the Oakley Mental Health Research Foundation, the University of Auckland Faculty Research Development fund, and Vodafone New Zealand Ltd. Funders were not involved in the conduct of the study or interpretation of the findings. The authors would like to acknowledge all of those involved in intervention development and the study including our team of research assistants, the schools and school guidance counselors, Shireen Chua, Karen Carter, Richard Smith, Philip Chao, Graham Burt, Victor Gaitan from EatLeadStudios, and The Hyperfactory.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEMO example video message of a New Zealand teen talking about their video diary.

[MPG File, 4MB - jmir_v14i1e13_app1.mpg]

Multimedia Appendix 2

CONSORT-EHEALTH (V1.6) checklist [45].

[PDF File (Adobe PDF File), 808KB - jmir_v14i1e13_app2.pdf]

References

1. 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 University Press; 1996.
2. 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]
3. Lewinsohn PM, Rohde P, Seeley JR. Major depressive disorder in older adolescents: prevalence, risk factors, and clinical implications. Clin Psychol Rev 1998 Nov;18(7):765-794. [Medline: 9827321]
4. Harrington R, Fudge H, Rutter M, Pickles A, Hill J. Adult outcomes of childhood and adolescent depression. I. Psychiatric status. Arch Gen Psychiatry 1990 May;47(5):465-473. [Medline: 2184797]
5. Kovacs M. Presentation and course of major depressive disorder during childhood and later years of the life span. J Am Acad Child Adolesc Psychiatry 1996 Jun;35(6):705-715. [doi: 10.1097/00004583-199606000-00010] [Medline: 8682751]
6. Brent DA, Birmaher B. Clinical practice. Adolescent depression. N Engl J Med 2002 Aug 29;347(9):667-671. [doi: 10.1056/NEJMcp012249] [Medline: 12200555]
7. Fleming JE, Boyle MH, Offord DR. The outcome of adolescent depression in the Ontario Child Health Study follow-up. J Am Acad Child Adolesc Psychiatry 1993 Jan;32(1):28-33. [doi: 10.1097/00004583-199301000-00005] [Medline: 8428881]
8. Adolescent Health Research Group. A health profile of New Zealand youth who attend secondary school. N Z Med J 2003 Apr 4;116(1171):U380. [Medline: 12740630]
9. Clark T. Te Ara Whakapiki Taitamariki: Youth '07: The Health and Wellbeing of Secondary School Students in New Zealand: Results for Maori Young People. Auckland, NZ: University of Auckland; 2008.
10. Adolescent Health Research Group. Te Ara Whakapipi Taitamariki: Maori Specific Findings of Youth2000. Auckland, NZ: The University of Auckland; 2004.
11. Rasanthan K, Ameratunga S, Chen J, Robinson E, Young W, Wong G, et al. A Health Profile of Young Asian New Zealanders Who Attend Secondary School: Findings From Youth2000. Auckland, NZ: University of Auckland; 2006.
12. Watanabe N, Hunot V, Omori IM, Churchill R, Furukawa TA. Psychotherapy for depression among children and adolescents: a systematic review. Acta Psychiatr Scand 2007 Aug;116(2):84-95. [doi: 10.1111/j.1600-0447.2007.01018.x] [Medline: 17650269]
13. 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] [Medline: 16435960]
14. Merry S, McDowell H, Hetrick S, Bir J, Muller N. Psychological and/or educational interventions for the prevention of depression in children and adolescents. Cochrane Database Syst Rev 2004(1):CD003380. [doi: 10.1002/14651858.CD003380.pub2] [Medline: 14974014]
15. Egger M, Smith GD, Altman D. Systematic Reviews in Health Care: Meta-analysis in Context. 2nd edition. London: BMJ Publishing Group; 2001.
16. Weisman SM, Graham DY. Evaluation of the benefits and risks of low-dose aspirin in the secondary prevention of cardiovascular and cerebrovascular events. Arch Intern Med 2002 Oct 28;162(19):2197-2202 [FREE Full text] [Medline: 12390062]
17. 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 [FREE Full text] [Medline: 11735841]
18. Merry S, McDowell H, Wild CJ, Bir J, Cunliffe R. A randomized placebo-controlled trial of a school-based depression prevention program. J Am Acad Child Adolesc Psychiatry 2004 May;43(5):538-547. [doi: 10.1097/00004583-200405000-00007] [Medline: 15100560]
19. Whittaker R, Borland R, Bullen C, Lin RB, McRobbie H, Rodgers A. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2009(4):CD006611. [doi: 10.1002/14651858.CD006611.pub2] [Medline: 19821377]
20. Mao Y, Zhang Y, Zhai S. Mobile phone text messaging for pharmaceutical care in a hospital in China. J Telemed Telecare 2008;14(8):410-414. [doi: 10.1258/jtt.2008.080406] [Medline: 19047450]
21. Strandbygaard U, Thomsen SF, Backer V. A daily SMS reminder increases adherence to asthma treatment: a three-month follow-up study. Respir Med 2010 Feb;104(2):166-171. [doi: 10.1016/j.rmed.2009.10.003] [Medline: 19854632]

22. Miloh T, Annunziato R, Arnon R, Warshaw J, Parkar S, Suchy FJ, et al. Improved adherence and outcomes for pediatric liver transplant recipients by using text messaging. Pediatrics 2009 Nov;124(5):e844-e850 [FREE Full text] [doi: 10.1542/peds.2009-0415] [Medline: 19822583]
23. Joo NS, Kim BT. Mobile phone short message service messaging for behaviour modification in a community-based weight control programme in Korea. J Telemed Telecare 2007;13(8):416-420. [doi: 10.1258/135763307783064331] [Medline: 18078554]
24. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
25. Haapala I, Barengo NC, Biggs S, Surakka L, Manninen P. Weight loss by mobile phone: a 1-year effectiveness study. Public Health Nutr 2009 Dec;12(12):2382-2391. [doi: 10.1017/S1368980009005230] [Medline: 19323865]
26. 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] [Medline: 19141433]
27. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
28. Tasker AP, Gibson L, Franklin V, Gregor P, Greene S. What is the frequency of symptomatic mild hypoglycemia in type 1 diabetes in the young?: assessment by novel mobile phone technology and computer-based interviewing. Pediatr Diabetes 2007 Feb;8(1):15-20. [doi: 10.1111/j.1399-5448.2006.00220.x] [Medline: 17341287]
29. Yoon KH, Kim HS. A short message service by cellular phone in type 2 diabetic patients for 12 months. Diabetes Res Clin Pract 2008 Feb;79(2):256-261. [doi: 10.1016/j.diabres.2007.09.007] [Medline: 17988756]
30. Kim HS, Kim NC, Ahn SH. Impact of a nurse short message service intervention for patients with diabetes. J Nurs Care Qual 2006;21(3):266-271. [Medline: 16816608]
31. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
32. Whittaker R, Dorey E, Bramley D, Bullen C, Denny S, Elley CR, et al. A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial. J Med Internet Res 2011;13(1):e10 [FREE Full text] [doi: 10.2196/jmir.1553] [Medline: 21371991]
33. Whittaker R, Maddison R, McRobbie H, Bullen C, Denny S, Dorey E, et al. A multimedia mobile phone-based youth smoking cessation intervention: findings from content development and piloting studies. J Med Internet Res 2008;10(5):e49 [FREE Full text] [doi: 10.2196/jmir.1007] [Medline: 19033148]
34. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
35. Meade CD, McKinney WP, Barnas GP. Educating patients with limited literacy skills: the effectiveness of printed and videotaped materials about colon cancer. Am J Public Health 1994 Jan;84(1):119-121. [Medline: 8279598]
36. Horne PJ, Tapper K, Lowe CF, Hardman CA, Jackson MC, Woolner J. Increasing children's fruit and vegetable consumption: a peer-modelling and rewards-based intervention. Eur J Clin Nutr 2004 Dec;58(12):1649-1660. [doi: 10.1038/sj.ejcn.1602024] [Medline: 15252421]
37. Robertson C, Gatchel RJ, Fowler C. Effectiveness of a videotaped behavioral intervention in reducing anxiety in emergency oral surgery patients. Behav Med 1991;17(2):77-85. [doi: 10.1080/08964289.1991.9935162] [Medline: 1878612]
38. Schunk DJ, Hanson AR, Cox PD. Peer-model attributes and children's achievement behaviors. J Educ Psychol 1987;79(1):54-61. [doi: 10.1037/0022-0663.79.1.54]
39. Scott LA, Black DR. Role modeling: an opportunity for the health education specialist. Health Educ Behav 1999 Oct;26(5):623-624. [Medline: 10533168]
40. Krouse HJ. Video modelling to educate patients. J Adv Nurs 2001 Mar;33(6):748-757. [Medline: 11298212]
41. 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]
42. Reynolds WM. Reynolds Adolescent Depression Scale, Second Edition (RADS-2): Professional Manual. Lutz, FL: Psychological Assessment Resources; 2002.
43. Poznanski EO, Mokros HB. Manual for the Children's Depression Rating Scale-Revised. Los Angeles, CA: Western Psychological Services; 1996.
44. Milne K, Bowler S, Li J, Salmon P. The Quit Group. 2009 Jul 31. Evaluation of the First Year of the Txt2Quit Service: 17 June 2008-16 June 2009 URL: http://www.quit.org.nz/file/research/FINAL%202008-09%20Txt2Quit%20evaluation%20report%2020090731.pdf [accessed 2011-11-08] [WebCite Cache ID 633YUJRl8]

45. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CBT: cognitive behavioral therapy
CDRS-R: Child Depression Rating Scale-Revised
NNT: number needed to treat
RADS-2: Reynolds Adolescent Depression Scale, Second Edition
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 08.06.11; peer-reviewed by M Kelley, S Kauer, W Evans; comments to author 13.09.11; revised version received 02.11.11; accepted 08.11.11; published 24.01.12.

Please cite as:
Whittaker R, Merry S, Stasiak K, McDowell H, Doherty I, Shepherd M, Dorey E, Parag V, Ameratunga S, Rodgers A
MEMO—A Mobile Phone Depression Prevention Intervention for Adolescents: Development Process and Postprogram Findings on Acceptability From a Randomized Controlled Trial
J Med Internet Res 2012;14(1):e13
URL: http://www.jmir.org/2012/1/e13/
doi:10.2196/jmir.1857
PMID:22278284

Original Paper

Perceptions and Experiences of Heart Failure Patients and Clinicians on the Use of Mobile Phone-Based Telemonitoring

Emily Seto[1,2], PhD, PEng; Kevin J Leonard[1,2], PhD, MBA; Joseph A Cafazzo[1,2,3], PhD, PEng; Jan Barnsley[2], PhD; Caterina Masino[1], MA; Heather J Ross[4,5], MD, MHSc, FRCPC

[1]Centre for Global eHealth Innovation, University Health Network, Toronto, ON, Canada

[2]Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

[4]Department of Medicine, University of Toronto, Toronto, ON, Canada

[5]Divisions of Cardiology and Transplant, University Health Network, Toronto, ON, Canada

Corresponding Author:
Emily Seto, PhD, PEng
Centre for Global eHealth Innovation
University Health Network
TGH/RFE Bldg, 4th Fl.
190 Elizabeth St
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4800 ext 6409
Fax: 1 416 340 3595
Email: emily.seto@uhn.on.ca

Abstract

Background: Previous trials of heart failure telemonitoring systems have produced inconsistent findings, largely due to diverse interventions and study designs.

Objectives: The objectives of this study are (1) to provide in-depth insight into the effects of telemonitoring on self-care and clinical management, and (2) to determine the features that enable successful heart failure telemonitoring.

Methods: Semi-structured interviews were conducted with 22 heart failure patients attending a heart function clinic who had used a mobile phone-based telemonitoring system for 6 months. The telemonitoring system required the patients to take daily weight and blood pressure readings, weekly single-lead ECGs, and to answer daily symptom questions on a mobile phone. Instructions were sent to the patient's mobile phone based on their physiological values. Alerts were also sent to a cardiologist's mobile phone, as required. All clinicians involved in the study were also interviewed post-trial (N = 5). The interviews were recorded, transcribed, and then analyzed using a conventional content analysis approach.

Results: The telemonitoring system improved patient self-care by instructing the patients in real-time how to appropriately modify their lifestyle behaviors. Patients felt more aware of their heart failure condition, less anxiety, and more empowered. Many were willing to partially fund the use of the system. The clinicians were able to manage their patients' heart failure conditions more effectively, because they had physiological data reported to them frequently to help in their decision-making (eg, for medication titration) and were alerted at the earliest sign of decompensation. Essential characteristics of the telemonitoring system that contributed to improved heart failure management included immediate self-care and clinical feedback (ie, teachable moments), how the system was easy and quick to use, and how the patients and clinicians perceived tangible benefits from telemonitoring. Some clinical concerns included ongoing costs of the telemonitoring system and increased clinical workload. A few patients did not want to be watched long-term while some were concerned they might become dependent on the system.

Conclusions: The success of a telemonitoring system is highly dependent on its features and design. The essential system characteristics identified in this study should be considered when developing telemonitoring solutions.

Key Words:

(J Med Internet Res 2012;14(1):e25) doi:10.2196/jmir.1912

KEYWORDS

heart failure; telemonitoring; mobile phone; patient monitoring; self-care; qualitative research

Introduction

Telemonitoring systems have been developed with the goal to improve outcomes and reduce the high costs associated with heart failure. However, the results from previous trials have been inconsistent, largely due to the diverse interventions under investigation and varying study designs [1-6]. There is currently a lack of insight into the features required for successful telemonitoring solutions.

A mobile phone-based telemonitoring system was developed with a user-centered design process, including iterative rounds of usability testing with heart failure patients and clinicians. The telemonitoring system was then evaluated through a randomized controlled trial (ClinicalTrials.gov NCT00778986) with 100 heart failure patients (n = 50 in each of the intervention and control groups). The primary intent of the trial was to pilot the telemonitoring system in order to determine the impact of the system on self-care, clinical management, and health outcomes. The quantitative findings from the trial suggested that the telemonitoring system improved quality of life through improved self-care and clinical management. The quantitative results are presented in detail in an accompanying paper published in this issue [7].

This paper discusses the qualitative findings from the trial based on in-depth patient and clinician interviews. The objective of the qualitative analysis was to obtain an understanding of the effects of the mobile phone-based telemonitoring system on self-care and clinical management. In addition, the analysis provides insight into the characteristics of a successful telemonitoring solution.

Methods

Study Design

The heart failure patient participants were recruited from the University Health Network (UHN) Heart Function Clinic in Toronto, Ontario, between September 2009 and February 2010. Semi-structured interviews were conducted with 22 heart failure patients who had used the telemonitoring system for 6 months. To be eligible for participation in this study, patients had to be older than 18 years of age, not be on the heart transplantation list, be expected to survive more than one year, and have a left ventricular ejection fraction (LVEF) < 40%. The UHN Research Ethics Board approved the study prior to commencement.

Although 50 heart failure patients used the telemonitoring system in total, saturation from the interview data was achieved after 22 patients. The choice of interviewees was generally based on who was first scheduled to visit the Heart Function Clinic, because the interviews were conducted face-to-face at their first scheduled clinic visit 6 months after recruitment. Efforts were made to interview patients who have had differing experiences with the telemonitoring system.

The five clinicians (three cardiologists and two nurse practitioners) from the Heart Function Clinic who interacted with the telemonitoring system during the study were also interviewed post-trial. Each patient and clinician semi-structured interview was between 15 and 60 minutes in duration. All interviews were audio recorded and later transcribed.

Study Intervention

The participants were provided with the telemonitoring system, in addition to standard care. They were asked to use the telemonitoring system to take daily morning weight and blood pressure readings, and to answer daily morning symptom questions (mainly yes/no questions) on a mobile phone for 6 months. The 17 patients who did not have an implantable cardioverter-defibrillator (ICD) were provided with an ECG recorder and asked to take weekly recordings (ECG recorder was not certified for use with ICDs). The weight and blood pressure readings (UA UC-321PBT weight scale and UA-767PBT blood pressure monitor, A&D Medical, USA) and ECG recordings (SelfCheck ECG PMP4, CardGuard, Israel) were automatically sent wirelessly via Bluetooth to a mobile phone (BlackBerry Pearl 8130, Research in Motion, Canada). A custom-designed and -built software application on the mobile phone was used to display and store the data, and to transmit the information to the data repository at the hospital.

Standard care at the UHN Heart Function Clinic included visits to the clinic from once every 2 weeks to once every 3-6 months, depending on the severity of the patient's heart failure condition and the need for optimization of their medication. Standard care also included heart failure education during preliminary visits at the Heart Function Clinic and the ability to telephone the clinic as necessary.

An instruction or alert was sent to the mobile phone based on all the physiological and symptom information. Both the patients and clinicians were able to view the data on a secure website. All data were also stored and accessible on the mobile phone. If a patient did not take all the required measurements by 10 AM each day, an automated adherence reminder phone call was sent to their home telephone. The use of the home telephone number as a contact point was deemed preferable, in case the patient was not near the system, the BlackBerry was in another room, the BlackBerry was turned off, etc. Patients in the telemonitoring group were given an individual training session on how to use the system and were provided with the telemonitoring equipment to take home during the recruitment session. Participants could also access technical support by phone throughout the study. Alerts with all the relevant information, including the ECG recording as an attachment, if available, were emailed to the physician's mobile phone, if measurements were outside the target range or symptoms were reported.

Data Analysis

The transcripts were analyzed using a conventional content analysis approach [8]. The study coordinator and a second

reviewer independently analyzed and coded the transcripts with the software program NVivo version 7 (QSR International, Doncaster, Victoria, Australia). Key themes were identified by both researchers and then discussed until agreement on the themes was reached. For the patient post-trial interview data, member checking was performed with six telemonitoring group participants who had agreed to participate in the post-trial interviews. Efforts were made to interview patients who had differing experiences with the telemonitoring system. The six participants were mailed a summary of the themes from the qualitative analysis and then called at home a week later to discuss their thoughts on the themes.

Results

The demographic and clinical characteristics of the 22 interviewed participants were representative of the patient population attending the UHN Heart Function Clinic, including an average age of 57 (SD 14) and 82% (n = 18) male. The following is a discussion of the themes found from the conventional content analysis.

Increased Self-Care

The telemonitoring system enabled patients to appropriately modify their lifestyle behaviors (eg, salt and fluid restrictions, diuretic dose, and exercise) at the first sign of decompensation. Both the automated instructions from the telemonitoring system and the clinician phone calls during times of apparent decompensation provided timely feedback and reinforcement on what was the most appropriate course of action, including taking extra diuretic medication. Thus, patients received the instructions during "teachable moments". A "teachable moment" is when the ability to learn a particular task is possible, because the timing is right [9]. These improvements in self-care were enabled through several factors related to the use of the telemonitoring system as described below.

Improved Awareness and Knowledge of Heart Failure Condition

Patients expressed becoming more aware of their heart failure condition and their own body, because they were taking their physiological measurements and symptoms daily, and their weight and blood pressure targets were brought to their attention daily. The automated feedback and clinician phone calls also alerted the patients when their health appeared to be worsening.

> *…I never thought of my sodium. (The telemonitoring system) made me look at other things (even though) I thought I was healthy. (The cardiologist) made me more aware of my sodium levels by my weight. My weight was fluctuating and she said, 'You're using too much salt. Your blood pressure is a little high, you know.' I'm even more aware now than I was before I started this program.* [Patient #16]

The patients were also more aware of the cause and effect relationship between their lifestyle choices and their health. In particular, they were able to correlate diet, fluid consumption, medication adherence, and exercise with changes in their weight, blood pressure, and symptoms.

> *…It's really taught me what the correlation is between salt intake and weight and water retention. An above normal sodium intake will show up immediately the next day as a weight gain and then as you clear that out of your system it goes back.* [Patient #2]

Increased Reassurance/Reduced Anxiety

Patients expressed feeling reassured that someone was watching over them. Some patients had substantial anxiety prior to using the telemonitoring system, especially those who were newly diagnosed or those who recently had an acute cardiac episode. Many patients referred to the telemonitoring system as a "security blanket" and it was "like almost having a doctor right beside you". Patients stated that after taking their daily measurements in the morning, they could go about their day without worry, if the telemonitoring system sent them a message confirming that everything was normal. The patients' informal caregivers (ie, family members) also felt reassured by the telemonitoring system.

> *…I was in the hospital five times last year, because of anxiety and my potassium getting out (of control). That did not happen since the study. So, it saves trips to the clinic for sure and it removes anxiety.* [Patient #33]

The patients also felt more reassured, because they were more connected to their healthcare team and their clinicians had more information about their condition. They believed they would not "fall through the cracks" as they did before using the telemonitoring system.

> *…It tends to eliminate one of the biggest problems of being sick and that's a sense of isolation, because I know that there's regular (ongoing) contact. So, if I'm not feeling well, I know I'm going to be getting a phone call and it seems to me that's worth gold.* [Patient #2]

Increased Empowerment and Confidence

Patients expressed feeling more in control, confident, and accountable, because they could directly observe the effects of their lifestyle choices on their health and become active participants in their own health. As noted above, many patients learned the correlation between consuming sodium and changes in weight and blood pressure. Some patients also received automated reminders to take extra diuretic medication after a weight gain, which confirmed taking the extra medication was the correct course of action. This group of patients received prior instruction from their cardiologist to take extra diuretic medication in this situation, but they still often felt uncertain of making the decision to take the extra medication on their own.

> *…I think I'm more responsible for myself, or more accountable I should say, because I know that I have to send this in and then I look at it. If my weight is higher, then I make sure that I make changes in the diet or make changes in the food restrictions that I'm supposed to and, if it's low, then I can also make (the) changes that I need.* [Patient #9]

Increased Self-Care Motivation

Patients expressed being more motivated to adhere to the recommended daily monitoring of weight, blood pressure, and symptoms owing to the telemonitoring system for several reasons. First, the patients knew their clinicians would find out, if they did not perform their daily measurements. This provided an incentive for the patients to perform the measurements, because they did not want their healthcare providers to think they were not following their instructions. Second, the patients expressed sensing that their measurements were now being used and interpreted. Prior to the trial, some of the patients stopped taking and recording their measurements, especially blood pressure readings, because they did not know how to interpret the values themselves and clinicians were not reviewing them.

> *…The fact that I get a call (about) my weight, that means somebody is really looking at this, so that was a bit of a confidence booster. The fact that this isn't just being chucked.* [Patient #13]

Third, the telemonitoring system helped them to establish a habit of taking measurements first thing in the morning.

> *…I kind of liked the fact that I was sort of following a routine and actually checking my weight and my blood pressure. It sort of gave me some comfort that I was sort of doing that kind of stuff. It's all those doubts that go through your mind. Having some routine sort of stabilizes myself in the morning.* [Patient #18]

Improved Clinical Management

The clinicians believed telemonitoring improved clinical management, because the system provided increased patient data for decision support and the automated alerts notified them at the earliest sign of decompensation.

Patient Data for Clinical Decision Support

The clinicians thought the telemonitoring system provided comprehensive information to support clinical decision-making. For example, the real-time daily weight, blood pressure, and symptom data were used to initiate 105 additional medication adjustments and instructions over the six-month trial. Of particular importance was a statistically significant increase in the number of patients in the telemonitoring group who were prescribed an aldosterone antagonist (type of diuretic medication) during the trial.

Previous studies have found less than a third of eligible patients receive heart failure guideline-recommended aldosterone antagonist therapy, partially because of the need to closely monitor serum potassium levels, due to the risk of hyperkalemia (a potentially fatal condition from elevated concentration of potassium in the blood) [10, 11]. The increased use of aldosterone antagonists in the telemonitoring group may reflect the ability of clinicians to closely follow their patients' daily weights and blood pressures, enhancing the feeling of security with medication up-titration. Patients could be easily instructed to see their primary care physician for follow-up of serum potassium levels as necessary. The involvement of the primary care physician was not excluded in the trial. All patients seen in the Heart Function Clinic have routine follow-up with their primary care physician or community cardiologist. Communication is maintained after all clinic visits and telephone communication is encouraged for issues or clinical concerns. The system could be easily adapted to allow monitoring by a nurse practitioner, primary care physician, internist, or community cardiologist.

> *…I think it was a very useful tool and that it has complemented the usual care that we provide to our patients. It was a way to detect things that we don't usually or we are unable to detect, because of ambulatory issues that some patients do not report. In many cases, it helped us make decisions, admissions, change of medications, closer follow up, and we have changed, a little bit, the (care) plan.* [Cardiologist #2]

Alerts of Patient Decompensation

The clinicians believed the telemonitoring system enabled them to detect decompensation earlier, because of the alerts generated from the daily patient physiological and symptom data. The clinicians considered the ability to individualize the system for different patients to be very important (ie, adjusting their target values). They also thought the system prompted them to interact with their patients at the most appropriate time in order to reinforce appropriate behaviors (eg, reduction of salt intake or extra diuretic medication after a sudden weight gain).

> *...Absolutely, it's a learning tool. At the beginning of this, they had no idea that their weight can change and they have a target. They learn, because they received the call: 'okay, you gained three pounds, take an extra lasix' [diuretic medication]. So, they know that, and in the future they will do it by themselves.* [Cardiologist #2]

Other Perceived Benefits

Both patients and clinicians were motivated to continue using the telemonitoring system, because of the perceived benefits. Besides the improved self-care and clinical management discussed above, there was a perception telemonitoring would reduce clinic visits and hospitalizations. The ease of use and portability of the system were also found to be benefits of the mobile phone-based telemonitoring system.

Reduction in Clinic Visits and Hospitalizations

Reducing the number of trips to the Heart Function Clinic was one of the most commonly cited potential benefits by patients. This was especially important for patients who lived in areas far from the clinic, felt very ill, and found coming to the clinic to be a financial burden. Although the trial was underpowered to detect changes in number of hospitalizations, clinicians believed telemonitoring could reduce the rate of hospitalizations, which could lead to cost savings. They mentioned that preventing one hospital admission per patient could pay for the cost of the telemonitoring system.

> *…The cost effectiveness (determined from) the number of hours and effort, and the cost of retaining the system (versus the savings from reduced*

hospitalization) is to be determined, but my sense is it definitely will save the department and the hospital (money). [Cardiologist #3]

Ease of Use

The telemonitoring system required the patients to take only a few minutes each morning to send in the physiological values. None of the participants had expressed issues using the system. Special consideration for those with vision impairment was accommodated in the design of the application, such that information was displayed with a large typeface and high contrast (black lettering on a white background, with no extraneous information). Even very elderly patients (oldest study participant was 88 years old) and those with no mobile phone experience were able to successfully use the telemonitoring system. Some patients required up to a week to become comfortable using the system or phoned for technical support, especially those who were less technologically savvy or who were not completely fluent in English. Many patients also had family members who could help them when necessary.

…That's a big advantage, because now they're not having to learn how to use a modem and anything else. (The mobile phone) gets a reading, if it seems like a valid reading, off it goes. [Patient #33]

Portability

Patients commented that the portability of the mobile phone-based telemonitoring system was useful, because it could be taken on vacation or to the cottage. The available website was seldom used by patients (only 13 patients ever logged onto the website), because all the relevant information was displayed on the mobile phone. Therefore, patient access to a computer was unnecessary.

...It gave me a lot of peace of mind, particularly (when) I was in Florida without a doctor nearby for three months, and it was just wonderful to know that my vital signs were being monitored. I'm very, very impressed at the responses that I got for those few times that the results were not within the guidelines. [Patient #18]

Clinicians liked the portability of the telemonitoring system, because they could monitor their patients from anywhere since the alerts were sent to their mobile phones. The website was also infrequently used by the clinicians, because the email alerts usually provided all the required information and enabled the clinician to simply click on the displayed patient's phone number to dial the number.

…I was at my cottage or when I was away on a weekend, I was always able to be (informed about my patients). I was in Bermuda and calling patients from Bermuda when I was down there on a four-day holiday weekend, and I would get a hold of the patients and stay on top of things. [Cardiologist #1]

Barriers to Long-Term Implementation

The large majority of patients wanted to continue to use the telemonitoring system indefinitely. The clinicians wanted to integrate the system into the Heart Function Clinic. However, there were several concerns raised regarding its long-term use.

Ongoing Costs

The clinicians and patients indicated cost was one of the main barriers to implementing the telemonitoring system on a long-term basis, although the cost to deliver this mobile phone service is projected to be much lower than conventional telemonitoring systems. In general, patients thought the healthcare system should pay for the use of the telemonitoring system through the Ontario Health Insurance Plan (OHIP), because it would reduce healthcare system costs in the long run. Patients were asked if they would be willing to pay for the telemonitoring system (ie, equipment and monthly cellular phone/data charges). Some patients were willing to partially pay for the use of the telemonitoring system, if it was a reasonable amount. In fact, some patients offered to pay at the end of the trial to be able to keep using the telemonitoring system.

…I think the healthcare system is going to save money. I don't know what the costs are of those devices, but I know that I have not been in the emergency department over night or for a period of time since we started this. Of course, everybody would like OHIP to pay for it. In my case, it's worth money to me and I would (pay for it). You'd be afraid not to do it. [Patient #33]

Some patients suggested having to pay a nominal fee might even encourage patients to adhere to taking the measurements.

...If I had to pay five dollars or ten dollars every month to be on this system, then I think it would keep me more consistent and saying, 'okay, I'm paying for this thing, I might as well do it'. They go, 'if it's not costing me anything, I don't care', whereas if it costs you something you tend to pay a little more attention to it. [Patient #18]

However, some patients indicated they did not have the financial means to help pay for using the telemonitoring system. Patients also suggested insurance companies or work benefits might be able to help fund the telemonitoring system. A few patients mentioned patients might be able to claim the use of the telemonitoring system as a medical expense on their income taxes, but it would be only a small return.

…Willing to pay is a hard question, as some people are on disability and cannot afford it. [Patient #34]

Participants were asked how much they would be willing to pay per month to continue using the telemonitoring system. Figure 1 displays the responses to this question. Fourteen patients responded they would not pay to use the telemonitoring system, but many of these patients stated in the comment section that they did not have the financial means to pay and the healthcare system should fund it. The second most common response was paying between Can $25 and Can $49 per month (8 respondents).

Figure 1. Responses to amount per month patients would be willing to pay to use the telemonitoring system.

Increased Clinical Work Load and Workflow Changes

One of the cardiologists' main concerns was the potential lack of time to respond to alerts, due to their busy schedules, if the telemonitoring system was permanently integrated into the clinic. They suggested a nurse practitioner manage the alerts, instead of a cardiologist.

> *…I think it's ideally suited for a nurse practitioner. Most of the issues you're dealing with are straightforward. They really probe into what a patient has been up to, and they kind of scratch below the surface more.* [Cardiologist #3]

The nurse practitioners were in favor of managing the alerts themselves. However, it was recognized that a nurse practitioner would need dedicated time to respond to the alerts and should become familiar with the patients on the telemonitoring system. The vast majority of the alerts would be generated in the morning after patients take their daily measurements. For nights and weekends, the cardiologist on call could respond to the alerts. The email alerts could be sent to a specific mobile phone that could be passed between the clinicians managing the alerts.

> *…I'd have no problems (managing the alerts). I think that would be great. I now communicate very closely with patients around their weight and lasix dosing and stuff like that. So, they call me directly and sometimes I'll talk to them daily. I find that as a result of that, they feel better connected.* [Nurse Practitioner #1]

Readiness of the Clinic for Integration

A challenge to integrating the telemonitoring system into the clinic would be maintaining up-to-date blood test values and medication, in order to provide current information in the email alerts to the clinicians. The hospital's electronic health record (EHR) was not always up-to-date (eg, outside blood test results and changes in medication between clinic visits were recorded on a paper chart). In addition, the data were usually in a free text clinical note that could not be easily used to populate a database. Therefore, changes in medication and blood test values were manually entered into a separate telemonitoring system database for the trial and manually updated monthly by a clinical research fellow. Ideally, the telemonitoring system and EHR should be integrated in the future and all relevant information should be available electronically.

> *…If someone had blood work done a week ago or even two weeks ago and you're thinking about adjusting their diuretic, well then you've got some useful information, (but not if) you have a potassium or creatinine that's three months old.* [Cardiologist #3]

Another issue that arose during the trial was some of the automated alerts advised the patient to phone the Heart Function Clinic. However, someone was not always available to answer the phone, and no one answered the phone outside of regular work hours.

Being Watched Long-term and System Dependency

A very small minority of patients expressed not wanting to use the telemonitoring system long-term, even though they thought there was a benefit to using it for a few months in order to learn more about self-care. They did not like "being watched" long-term, because they wanted "to enjoy life once in a while".

> *…I feel like a prisoner. It doesn't take much to gain three pounds. I used to drink three or four beers. Next day, I know my weight is going to be higher. I don't*

> *want a phone call, you know. It's kind of pressure... it's pressuring me, you know.* [Patient #35]

Another potential negative effect of using the telemonitoring system is that some patients may feel dependent on it.

> *…(Taking the monitoring system away is like) another crutch that you're losing. You become conditioned to having it there as a backup, and there would be an initial sort of sense of loss.* [Patient #29]

Discussion

Our study provided an in-depth investigation into the perceptions and experiences of heart failure patients and clinicians on the use of a mobile phone-based telemonitoring system and the mechanisms that helped improve self-care and clinical management. For example, self-care improved, because the patients felt more aware and knowledgeable regarding their heart failure condition, less anxiety, and more empowered and motivated to improve their condition. However, both the clinicians and patients had some concerns on using the telemonitoring system long-term, such as obtaining operational funding for the telemonitoring system and potentially increased clinical workload. This study also provided insight into system characteristics that are essential for the success of telemonitoring on heart failure outcomes. In particular, these important characteristics are (1) immediate self-care feedback, (2) immediate clinical feedback as necessary, (3) ease of use, and (4) perceived benefits of the telemonitoring system for continued adherence.

Immediate Self-Care Feedback

Appropriate heart failure self-care has been found to improve health outcomes and reduce healthcare costs [12-14]. Self-care has also been stressed as an important component of heart failure management in international clinical guidelines [15-17]. Our trial indicated the use of the telemonitoring system improved self-care as measured through the Self-Care of Heart Failure Index (SCHFI) [18]. In particular, a comparison of the post-trial data between the telemonitoring and control groups found a statistically significant difference in SCHFI maintenance scores (P= .03), indicating the telemonitoring group had greater self-care maintenance (ie, a higher SCHFI maintenance score). From the patient interviews, it was evident they were able to modify their own lifestyle behaviors from the automated immediate feedback of the telemonitoring system and the timely self-care feedback from their clinicians.

The lack of real-time, self-care feedback may be one of the primary reasons previous telemonitoring trials, such as the Telemonitoring to Improve Heart Failure Outcomes (Tele-HF) Trial, have failed to show positive health outcomes [3]. The Tele-HF trial intervention was an interactive voice-response system the patient called to record heart failure symptoms and weight data. The trial found no reduction in mortality or hospital admissions, but the system was lacking real-time patient feedback on self-care instructions and alerts.

Immediate Clinical Feedback

From the clinicians' perspective, the telemonitoring system alerted them to contact their patients, such as for continuous optimization of medications. From the patients' perspective, the immediate clinical feedback not only improved health outcomes, but also reduced their anxiety by knowing that someone was watching over them and motivated them to improve self-care. In the Tele-HF trial, site coordinators reviewed patient information daily during weekdays and should have contacted the patient as required. It is thus possible the response by clinicians to indications of decompensation was not timely.

Ease of Use

Both patients and clinicians found the telemonitoring system to be easy to use. The system was highly automated and required minimal understanding of technology, including mobile phones. Previous trials of heart failure telemonitoring have demonstrated that when patients find the system difficult to use, adherence to taking the physiological measurements is poor. The MOBIle TELemonitoring in Heart Failure Study (MOBITEL) asked the intervention group (n = 54) to send their daily measurements of blood pressure, heart rate, body weight, and dosage of heart failure medication via a mobile phone's Internet browser [19]. Due to the patients having difficulty in using the browser to transmit the data, even after intensive training, 12 patients dropped out immediately.

Perceived Benefits

Participants in the current trial completed their daily measurements on average 5 to 6 days per week throughout the six-month duration. The high adherence to daily measurements can be partially attributed to the ease of use of the system. However, patients must also believe that transmitting daily information will benefit them in tangible ways. During the pre-trial patient interviews, many of the patients stated they stopped taking their blood pressure at home, because the readings were not sent to their clinicians and they did not know how to interpret or act on the blood pressure readings on their own [20].

The clinicians supported the use of telemonitoring, because they thought it was a useful tool to complement the management of their patients' heart failure conditions and to help increase self-care. They believed telemonitoring would ultimately improve their patients' health outcomes, including reducing the number of hospital admissions. Although there were concerns of implementing the telemonitoring system, such as increased workload, the clinicians were willing to try to resolve any barriers, because of the numerous perceived benefits from mobile phone-based telemonitoring.

Previous Studies on Perceptions of Telemonitoring

Our study supported some findings from previous studies on the thoughts of patients and clinicians on telemonitoring for various chronic illnesses. For example, previous studies have found that ease of use, perceived tangible benefits, and cost-effectiveness are important aspects of mobile phone-based telemonitoring adoption for heart failure, asthma, and blood pressure management [20-23]. However, whether telemonitoring increases or decreases patient anxiety appears to be dependent on the particular patient population. A study on nocturnal home hemodialysis found anxiety reduced with telemonitoring,

especially for the patient's caregiver [24], which is similar to the findings from the present trial. Conversely, a study on blood pressure telemonitoring found that a major concern was increased patient anxiety, such as from a single high blood pressure reading [23]. Possibly, as the perception of the severity of the chronic illness or event being monitored increases, the more beneficial telemonitoring is at reducing anxiety.

Limitations

A limitation to the findings from this study is their general applicability, because of the specific group of heart failure patients chosen to participate. The participants were all recruited from the Heart Function Clinic and, therefore, may have been more motivated to adhere to the telemonitoring protocol than patients followed by primary physicians. The participants were also younger on average than the average heart failure patient, because the UHN Heart Function Clinic treats a high proportion of severely ill patients, including some very young patients (ie, in their twenties). However, many of the study participants were elderly. In addition, many of the participants had stable heart failure (ie, had not been admitted into hospital for several years). Higher-risk patients, who are frequently readmitted to hospital, might benefit even more from telemonitoring. Further investigation into the characteristics of heart failure patients, who would be suitable for and benefit from mobile phone-based telemonitoring, is required [4].

Finally, clinicians who participated in the study may have been more willing to use the telemonitoring system than primary physicians and those at other heart function clinics. Prior to the trial, some of the clinicians used their own personal time to communicate with their patients and kept in close contact with their high-risk patients through various means (ie, telephone, email, and fax). Therefore, despite their concerns about potentially increased workload, they may have been more amenable to the idea of mobile phone-based telemonitoring.

Conclusions

Mobile phone-based telemonitoring enabled patients to positively change their lifestyle behaviors and to improve their quality of life, including reducing anxiety. The telemonitoring system also improved clinical management, by providing real-time physiological information and alerts. The in-depth qualitative analysis revealed several system characteristics that contributed to improved heart failure management, such as (1) having immediate self-care and clinical feedback, (2) being easy and quick to use, and (3) providing tangible benefits to the end-users (ie, the patients). These system characteristics will help ensure patient adherence and clinician buy-in, which are both necessary for successful implementation of any telemonitoring solution.

Acknowledgments

Funding for this work was in part provided by the Toronto General Hospital Foundation and the Natural Sciences and Engineering Research Council of Canada Strategic Research Network Grant entitled Healthcare Support through Information Technology Enhancements (hSITE).

Conflicts of Interest

None declared.

References

1. Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K, et al. Home telehealth for chronic obstructive pulmonary disease: a systematic review and meta-analysis. J Telemed Telecare 2010;16(3):120-127. [doi: 10.1258/jtt.2009.090812] [Medline: 20197355]
2. Inglis SC, Clark RA, McAlister FA, Ball J, Lewinter C, Cullington D, et al. Structured telephone support or telemonitoring programmes for patients with chronic heart failure. Cochrane Database Syst Rev 2010(8):CD007228. [doi: 10.1002/14651858.CD007228.pub2] [Medline: 20687083]
3. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
4. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, et al; Telemedical Interventional Monitoring in Heart Failure Investigators. Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure: the telemedical interventional monitoring in heart failure study. Circulation 2011 May 3;123(17):1873-1880. [doi: 10.1161/CIRCULATIONAHA.111.018473] [Medline: 21444883]
5. Clark RA, Inglis SC, McAlister FA, Cleland JG, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. BMJ 2007 May 5;334(7600):942 [FREE Full text] [doi: 10.1136/bmj.39156.536968.55] [Medline: 17426062]
6. Klersy C, De Silvestri A, Gabutti G, Regoli F, Auricchio A. A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 2009 Oct 27;54(18):1683-1694. [doi: 10.1016/j.jacc.2009.08.017] [Medline: 19850208]
7. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile Phone-Based Telemonitoring for Heart Failure Management: A Randomized Controlled Trial. J Med Internet Res 2012 (forthcoming). [doi: 10.2196/jmir.1909]
8. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

9. Havighurst RJ. Human Development and Education. New York: Longmans, Green; 1953.
10. Albert NM, Yancy CW, Liang L, Zhao X, Hernandez AF, Peterson ED, et al. Use of aldosterone antagonists in heart failure. JAMA 2009 Oct 21;302(15):1658-1665. [doi: 10.1001/jama.2009.1493] [Medline: 19843900]
11. Zannad F, McMurray JJ, Krum H, van Veldhuisen DJ, Swedberg K, Shi H, et al; EMPHASIS-HF Study Group. Eplerenone in patients with systolic heart failure and mild symptoms. N Engl J Med 2011 Jan 6;364(1):11-21 [FREE Full text] [doi: 10.1056/NEJMoa1009492] [Medline: 21073363]
12. Vinson JM, Rich MW, Sperry JC, Shah AS, McNamara T. Early readmission of elderly patients with congestive heart failure. J Am Geriatr Soc 1990 Dec;38(12):1290-1295. [Medline: 2254567]
13. McAlister FA, Stewart S, Ferrua S, McMurray JJ. Multidisciplinary strategies for the management of heart failure patients at high risk for admission: a systematic review of randomized trials. J Am Coll Cardiol 2004 Aug 18;44(4):810-819. [doi: 10.1016/j.jacc.2004.05.055] [Medline: 15312864]
14. Krumholz HM, Amatruda J, Smith GL, Mattera JA, Roumanis SA, Radford MJ, et al. Randomized trial of an education and support intervention to prevent readmission of patients with heart failure. J Am Coll Cardiol 2002 Jan 2;39(1):83-89. [Medline: 11755291]
15. Swedberg K, Cleland J, Dargie H, Drexler H, Follath F, Komajda M, et al; Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology. Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology. Eur Heart J 2005 Jun;26(11):1115-1140 [FREE Full text] [doi: 10.1093/eurheartj/ehi204] [Medline: 15901669]
16. Heart Failure Society Of America. HFSA 2006 Comprehensive Heart Failure Practice Guideline. J Card Fail 2006 Feb;12(1):e1-e2. [doi: 10.1016/j.cardfail.2005.11.005] [Medline: 16500560]
17. Riegel B, Moser DK, Anker SD, Appel LJ, Dunbar SB, Grady KL, et al; American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Clinical Cardiology, American Heart Association Council on Nutrition, Physical Activity, and Metabolism, American Heart Association Interdisciplinary Council on Quality of Care and Outcomes Research. State of the science: promoting self-care in persons with heart failure: a scientific statement from the American Heart Association. Circulation 2009 Sep 22;120(12):1141-1163 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192628] [Medline: 19720935]
18. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the self-care of heart failure index. J Cardiovasc Nurs 2009;24(6):485-497. [doi: 10.1097/JCN.0b013e3181b4baa0] [Medline: 19786884]
19. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, et al; MOBITEL Investigators. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009;11(3):e34 [FREE Full text] [doi: 10.2196/jmir.1252] [Medline: 19687005]
20. Seto E, Leonard KJ, Masino C, Cafazzo JA, Barnsley J, Ross HJ. Attitudes of heart failure patients and health care providers towards mobile phone-based remote monitoring. J Med Internet Res 2010;12(4):e55 [FREE Full text] [doi: 10.2196/jmir.1627] [Medline: 21115435]
21. Cleland J, Caldow J, Ryan D. A qualitative study of the attitudes of patients and staff to the use of mobile phone technology for recording and gathering asthma data. J Telemed Telecare 2007;13(2):85-89. [doi: 10.1258/135763307780096230] [Medline: 17359572]
22. Pinnock H, Slack R, Pagliari C, Price D, Sheikh A. Professional and patient attitudes to using mobile phone technology to monitor asthma: questionnaire survey. Prim Care Respir J 2006 Aug;15(4):237-245 [FREE Full text] [doi: 10.1016/j.pcrj.2006.03.001] [Medline: 16843066]
23. Bostock Y, Hanley J, Douglas M, Pinnock H, Padfield P, McKinstry B. The acceptability to patients and professionals of remote blood pressure monitoring using mobile phones. Primary Health Care Research & Development 2009;10:299-308. [doi: 10.1017/S1463423609990107]
24. Cafazzo JA, Leonard K, Easty AC, Rossos PG, Chan CT. Patient perceptions of remote monitoring for nocturnal home hemodialysis. Hemodial Int 2010 Oct;14(4):471-477. [doi: 10.1111/j.1542-4758.2010.00473.x] [Medline: 20955280]

Edited by H Potts; submitted 16.08.11; peer-reviewed by B McKinstry, D Patterson; comments to author 05.09.11; revised version received 04.12.11; accepted 12.12.11; published 10.02.12.

Please cite as:
Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ
Perceptions and Experiences of Heart Failure Patients and Clinicians on the Use of Mobile Phone-Based Telemonitoring
J Med Internet Res 2012;14(1):e25
URL: http://www.jmir.org/2012/1/e25/
doi:10.2196/jmir.1912
PMID:22328237

Original Paper

Mobile Phone-Based Telemonitoring for Heart Failure Management: A Randomized Controlled Trial

Emily Seto[1,2], PhD, PEng; Kevin J Leonard[1,2], PhD, MBA; Joseph A Cafazzo[1,2,3], PhD, PEng; Jan Barnsley[2], PhD; Caterina Masino[1], MA; Heather J Ross[4,5], MD, MHSc, FRCPC

[1]Centre for Global eHealth Innovation, University Health Network, Toronto, ON, Canada

[2]Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

[4]Department of Medicine, University of Toronto, Toronto, ON, Canada

[5]Divisions of Cardiology and Transplant, University Health Network, Toronto, ON, Canada

Corresponding Author:
Emily Seto, PhD, PEng
Centre for Global eHealth Innovation
University Health Network
TGH/RFE Bldg, 4th Fl.
190 Elizabeth St.
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4800 ext 6409
Fax: 1 416 340 3595
Email: emily.seto@uhn.on.ca

Abstract

Background: Previous trials of telemonitoring for heart failure management have reported inconsistent results, largely due to diverse intervention and study designs. Mobile phones are becoming ubiquitous and economical, but the feasibility and efficacy of a mobile phone-based telemonitoring system have not been determined.

Objective: The objective of this trial was to investigate the effects of a mobile phone-based telemonitoring system on heart failure management and outcomes.

Methods: One hundred patients were recruited from a heart function clinic and randomized into telemonitoring and control groups. The telemonitoring group (N = 50) took daily weight and blood pressure readings and weekly single-lead ECGs, and answered daily symptom questions on a mobile phone over 6 months. Readings were automatically transmitted wirelessly to the mobile phone and then to data servers. Instructions were sent to the patients' mobile phones and alerts to a cardiologist's mobile phone as required.

Results: Baseline questionnaires were completed and returned by 94 patients, and 84 patients returned post-study questionnaires. About 70% of telemonitoring patients completed at least 80% of their possible daily readings. The change in quality of life from baseline to post-study, as measured with the Minnesota Living with Heart Failure Questionnaire, was significantly greater for the telemonitoring group compared to the control group (P = .05). A between-group analysis also found greater post-study self-care maintenance (measured with the Self-Care of Heart Failure Index) for the telemonitoring group (P = .03). Brain natriuretic peptide (BNP) levels, self-care management, and left ventricular ejection fraction (LVEF) improved significantly for both groups from baseline to post-study, but did not show a between-group difference. However, a subgroup within-group analysis using the data from the 63 patients who had attended the heart function clinic for more than 6 months revealed the telemonitoring group had significant improvements from baseline to post-study in BNP (decreased by 150 pg/mL, P = .02), LVEF (increased by 7.4%, P = .005) and self-care maintenance (increased by 7 points, P = .05) and management (increased by 14 points, P = .03), while the control group did not. No differences were found between the telemonitoring and control groups in terms of hospitalization, mortality, or emergency department visits, but the trial was underpowered to detect differences in these metrics.

Conclusions: Our findings provide evidence of improved quality of life through improved self-care and clinical management from a mobile phone-based telemonitoring system. The use of the mobile phone-based system had high adherence and was feasible for patients, including the elderly and those with no experience with mobile phones.

Trial Registration: ClinicalTrials.gov NCT00778986

(J Med Internet Res 2012;14(1):e31) doi:10.2196/jmir.1909

KEYWORDS

heart failure; telemedicine; mobile phone; patient monitoring; randomized controlled trial

Introduction

The demand for health care resources to manage heart failure is increasing with the aging population. Innovative methods to help alleviate this burden and to improve the poor outcomes from heart failure are required. Previous studies on traditional telemonitoring of heart failure patients (ie, using dedicated hardware as an information and transmission hub) have determined telemonitoring has the potential to reduce mortality, hospitalizations, and costs as well as improve quality of life, self-care, and New York Heart Association (NYHA) class [1-4]. However, the results of telemonitoring trials have been inconsistent largely due to diverse study interventions and variations in study design.

Little is known about the feasibility and effects of mobile phone-based telemonitoring systems. Investigating mobile systems is a logical next step because they enable greater scalability of telemonitoring due to their relatively low cost compared to traditional systems, and because they provide more freedom for the patient due to their portability. Two recent randomized controlled trials of mobile phone-based telemonitoring for heart failure management have been reported in literature [5,6]. The Telemedical Interventional Monitoring in Heart Failure (TIM-HF) trial by Koehler et al (2011) found no reductions in hospitalizations or mortality [5]. The authors concluded their study does not rule out the potential benefits of telemonitoring, but instead that there is a need to identify the heart failure population that could benefit from telemonitoring [5]. The trial by Scherr et al (2009), which required patients to enter readings using a mobile phone's Internet browser, highlighted the importance of system design on the success of the telemonitoring system [6]. Many patients found using the Internet browser to be too difficult, resulting in 12 out of the 54 patients in the intervention group being "never beginners."

The objective of our randomized controlled trial was to perform an in-depth investigation of the effects of a highly automated and user-centered mobile phone-based telemonitoring system on self-care and clinical management, with the aim of improving heart failure outcomes (Trial Registration: ClinicalTrials.gov NCT00778986). We used an extensive user-centered design process to develop the telemonitoring system in order to ensure it was as easy to use as possible and to meet the needs of the clinicians and patients. User-centered design refers to a philosophy that bases the design on information about and input from the people who will be using the product. The user-centered design process will be discussed in a separate publication.

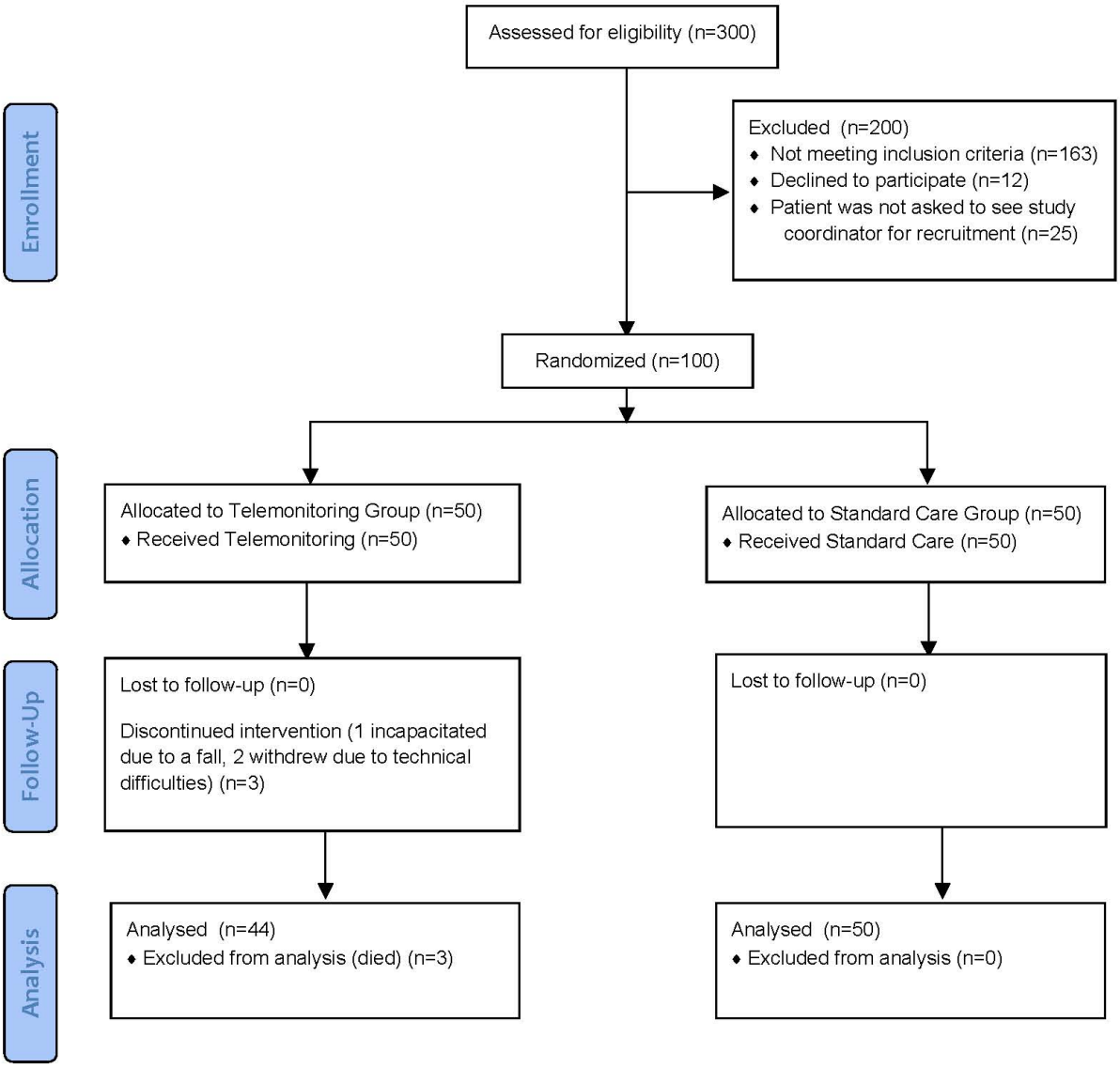
Methods

Study Participants

One hundred participants were recruited from the University Health Network (UHN) Heart Function Clinic in Toronto, Ontario, between September 2009 and February 2010 (Figure 1). The UHN Research Ethics Board approved the trial prior to commencement. The primary intent of the trial was to pilot the telemonitoring system in order to determine the impact of the system on self-care and clinical management. A sample size calculation was based on the Self-Care of Heart Failure Index (SCHFI), using a population standard deviation of 20 and an effect size of 10 (effect size represents a clinically significant change of more than half a standard deviation) as determined in previous studies (alpha = 0.05, power = 0.8) [7,8]. We calculated the required sample size per group to be 34, and recruited 50 participants for the intervention group and 50 for the control group to compensate for the patients estimated as lost to follow-up, including due to mortality, over the six-month trial.

Eligible participants were ambulatory patients diagnosed with heart failure. Other eligibility criteria included 18 years of age or older, ability to speak and read in English, not on the heart transplantation list, an expected survival of greater than one year, and a left ventricular ejection fraction (LVEF) less than 40%. During their Heart Function Clinic visit, patients who met the inclusion criteria (as deemed by their cardiologist), were invited to speak to the study coordinator (ES) regarding participation in the study. Each participant provided informed consent and received Can $24 as reimbursement for travel and parking expenses.

Figure 1. Flow of participants through the trial.



Study Protocol

The 100 participants were randomized into the telemonitoring (TM) group and standard care (SC) group using stratified four-block randomization. Stratification was based on NYHA classification (NYHA class II-III and NYHA class IV). There were no participants in NYHA class I. An online computer-generated randomization tool, Research Randomizer [9], was used to determine the order of participants in the telemonitoring and standard care groups. The study coordinator

was blinded to which group the patient would be assigned until each patient consented to participate in the trial.

Each patient received a questionnaire to complete at home. The questionnaire included demographic and clinical characteristic questions, and the SCHFI and the Minnesota Living with Heart Failure Questionnaire (MLHFQ), which are validated tools for measuring self-care and quality of life, respectively [10-12]. The SCHFI is made up of three subscales: self-care maintenance (choice of behaviors used to maintain physiological stability), self-care management (response to symptoms when they occur), and self-care confidence. The maintenance, management, and confidence subscales consist of 5, 6, and 4 Likert questions, respectively. A higher score on the SCHFI indicates improved self-care. The MLHFQ consists of 21 questions that use a 6-point Likert scale. The physical dimension score for the MLHFQ is the summation of 8 questions (eg, Did your heart failure make you sit or lie down to rest during the day?), while the emotional dimension score is the summation of 5 other questions (eg, Did your heart failure make you worry?). A lower score on the MLHFQ indicates higher quality of life.

The standard care group received standard care at the UHN Heart Function Clinic, which includes visiting the clinic between once every 2 weeks to once every 3 to 6 months, depending on the severity of the patient's heart failure condition and the need for optimizing their medication. Standard care also includes heart failure education during preliminary visits at the Heart Function Clinic and the ability to telephone the clinic as necessary. Participants in the standard care group were not contacted again regarding the study until the end of the trial.

The participants in the telemonitoring group received the telemonitoring system in addition to standard care. They were asked to use the telemonitoring system for 6 months to take daily morning weight and blood pressure readings as well as weekly single-lead electrocardiograms (ECGs) if provided with an ECG recorder. They were also asked to answer daily morning symptom questions on a mobile phone. Only the 17 patients who did not have an implantable cardioverter defibrillator (ICD) were provided with an ECG recorder because the recorder was not certified for use with ICDs. Patients were also told to report their symptoms through the mobile phone if they did not feel well during the day. The patients in the telemonitoring group were given an individual training session on how to use the system during the recruitment session, and were provided with technical support by telephone throughout the study. The daily measurements took about 5 minutes each morning.

Six months after recruitment, all participants were mailed a post-study questionnaire. Semi-structured interviews were conducted with participants in the telemonitoring group to determine their experiences with the system. Twenty-two patients were interviewed, at which point saturation was achieved (ie, no new themes were identified). The individual interviews lasted between 15 and 60 minutes and took place in a private consultation room at the Heart Function Clinic.

The five clinicians (three cardiologists and two nurse practitioners) from the Heart Function Clinic who managed the alerts and/or used the data from the telemonitoring system during the trial were also interviewed post-study. Each semi-structured interview lasted between 15 and 45 minutes. All patient and clinician interviews were audio-recorded and later transcribed.

Telemonitoring System Overview

The weight and blood pressure readings (UA UC-321PBT weight scale and UA-767PBT blood pressure monitor, A&D Medical, USA) and ECG recordings (SelfCheck ECG PMP4, CardGuard, Israel) were automatically sent wirelessly via Bluetooth to a mobile phone (BlackBerry Pearl 8130, Research in Motion, Canada) and then to the data repository at the hospital. Patients also answered symptom questions (mainly yes/no) through the mobile phones. The measurements sent to the data repository included a study identifier. No patient names were transmitted and no identifying information was stored on the mobile phone. The mobile phone displayed an instruction on what to do after taking each measurement. A final message or alert based on the physiological and symptom information was sent to the mobile phone (Figure 2). Clinicians were able to modify any necessary physiological target ranges per patient through a secure website. The alerts ranged from low priority ones (eg, to retake the measurements if the patient feels worse) to high priority alerts (eg, to go to the local emergency department or call 911). The patients and the clinicians were able to view the physiological data on a secure website in tabular and graphic formats. All the data were also stored and accessible on the mobile phone. If a patient did not take all the required measurements by 10 am each morning, an automated adherence reminder phone call was sent to their home telephone.

If measurements were outside of the target range or if the patient reported symptoms, alerts were emailed to the cardiologist's mobile phone. The email alerts included the patient's contact information, medication list, symptom and physiological information (including the ECG recording as an attachment, if available), latest serum creatinine and potassium, and the alert message sent to the patient. If the cardiologist determined contacting the patient was warranted, they were able to call the patient by selecting their phone number in the email. The patient alert message also instructed them to contact the Heart Function Clinic if they felt they should. During the trial, the cardiologist usually called the patient within a few minutes of receiving the alert. The cardiologist emailed the study coordinator specifying the action resulting from each alert (eg, calling the patient, modificaton of medications, etc) by replying to the original email alert.

Figure 2. Sample message sent to patient's mobile phone.

Outcome Measures

The primary outcomes of this study included a surrogate for heart failure prognosis, specifically brain natriuretic peptide (BNP), self-care as measured by the SCHFI, and quality of life as measured by the MLHFQ. Hospital readmissions, number of nights in hospital, and mortality were secondary outcome measures because the study was underpowered to detect differences between groups for these metrics. Other secondary outcome measures included number of emergency department visits and number of Heart Function Clinic visits. In addition, LVEF, NYHA class, medication prescriptions, and blood test results (specifically creatinine, sodium, potassium, hemoglobin, and urate values) were also subsequently analyzed.

Data Analysis

The normality of the data for each outcome measure was determined through Kolmogorov–Smirnov and Shapiro–Wilk tests of normality. Data that were normally distributed were MLHFQ, SCHFI maintenance, SCHFI management, sodium, potassium, hemoglobin, urate, and LVEF values. All other parameters were analyzed with non-parametric tests.

Between-group analyses using independent Student *t* tests and Mann–Whitney tests (for normally and not normally distributed data, respectively) were first performed to compare the telemonitoring group and standard care group post-study data. Between-group analyses were also performed to compare the change scores. Paired Student *t* tests and Wilcoxon signed rank tests were then performed to compare baseline and post-study

data within the telemonitoring and standard care groups. The statistical analyses were performed using the statistical software application SPSS 17.0 (IBM Corporation, USA). Statistical significance was considered at $P < .05$ unless otherwise specified. All test results reported are 2-tailed.

Interview data were analyzed using a conventional content analysis approach [13]. Two researchers (ES and CM) analyzed the transcripts independently and coded the transcripts with the software program NVivo version 7 (QSR International, Doncaster, Victoria, Australia). The researchers then discussed the themes and issues that emerged until a consensus was reached. The qualitative results are presented in detail in an accompanying paper [14].

Results

Twelve out of 112 patients approached to participate in this study declined. One patient felt overwhelmed by the idea of participating in the trial, two did not want to take measurements every day, one said he did not have the time, and one thought as a result of a stroke he would not be able to understand how to perform the monitoring. The remainder did not provide specific reasons for declining. The three patients from the telemonitoring group who withdrew from the study included a patient who became incapacitated during the trial due to a fall, and two participants who decided to withdraw from the study when they had technical difficulties with the telemonitoring equipment. No patients in the standard care group officially withdrew from the study (Figure 1).

Baseline Patient Data

Table 1 shows the baseline demographic and clinical characteristics of the 100 patients who participated in the trial. The profiles of the telemonitoring and standard care groups were similar and representative of the patient population attending the UHN Heart Function Clinic. A comparison of baseline study parameters showed no statistical differences between telemonitoring and standard care groups for any outcome measures. Baseline questionnaires were completed and returned by 94 patients (46 from the telemonitoring group and 48 from the standard care group) and 84 patients returned post-study questionnaires (39 from the telemonitoring group and 45 from the standard care group), while 82 patients returned both baseline and post-study questionnaires.

Telemonitoring System Utilization

Patient Adherence

Patients completed their required measurements on average between 5 to 6 days per week throughout the six-month trial (Figure 3). Adherence decreased only slightly from the beginning of the trial to the end. Adherence during the first week was relatively low because some patients had to travel for a number of days to get home before using the system, and some patients required technical telephone support before properly using the system. About 42, 33, and 16 out of the 50 telemonitoring group patients (84%, 66%, and 32%) completed at least 91 (50%), 146 (80%), and 173 (95%) of possible daily readings over the six months, respectively. Missed measurements were sometimes due to technical issues with the telemonitoring system or else because of patients going on vacation without bringing the monitoring equipment with them.

The adherence data presented are an underestimate of the true adherence because they are based on the adherence phone calls sent at 10 am if patients had not yet completed their daily measurements. Occasionally, patients would take their measurements after 10 am when the adherence reminder was already sent because they had woken after 10 am.

Clinical Utilization of the System

Table 2 summarizes the clinical utilization of the system and the actions the clinicians performed based on the alerts. One cardiologist received the majority (1367) of the email alerts. Another cardiologist received alerts (311 alerts) when covering for the primary cardiologist for a three-week period. Low priority alerts were more frequently generated than the more urgent alerts. The cardiologists (and occasionally a nurse practitioner) called patients 480 times over the 6 months, often to provide instructions or to educate. The most common clinical action was an instruction or change in medication (105 times). Other actions included ordering additional blood work, moving the patient's clinic visit forward, or instructing the patient to see their family physician or to go to the emergency department.

Table 1. Baseline demographic and clinical characteristics of patient participants[a]

Characteristic		telemonitoring group (N = 50)	standard care group (N = 50)
Age, mean (SD), years		55.1 (13.7)	52.3 (13.7)
Gender, No. (%)	Male	41 (82)	38 (76)
	Female	9 (18)	12 (24)
Ethnicity, No. (%)	Caucasian	39 (78)	33 (66)
	African Canadian	5 (10)	4 (8)
	South East Asian	2 (4)	2 (4)
	Chinese	0	3 (6)
	Other	4 (8)	8 (16)
Marital status, No. (%)	Married	34 (68)	28 (56)
	Not Married	12 (24)	19 (38)
Living arrangement, No. (%)	Living with partner or family	42 (84)	38 (76)
	Living alone	6 (12)	9 (18)
Highest education level achieved, No. (%)	Less than high school	1 (2)	6 (12)
	High school	12 (24)	13 (26)
	College/University	33 (66)	28 (56)
Income, No. (%)	< $15,000	9 (18)	11 (22)
	$15,000-$49,999	16 (32)	18 (36)
	> $50,000	15 (30)	13 (26)
	Preferred not to answer	6 (12)	5 (10)
Employment, No. (%)	Full-time	13 (26)	14 (28)
	Part-time	3 (6)	1 (2)
	Disability due to heart failure	18 (36)	19 (38)
	Retired	8 (16)	7 (14)
	Unemployed	4 (8)	7 (14)
NYHA class, No. (%)	II	21 (42)	22 (44)
	II/III	6 (12)	5 (10)
	III	21 (42)	21 (42)
	IV	2 (4)	2 (4)
Left ventricular ejection fraction, mean (SD)		27.1 (7.8)	27.0 (9.9)
Blood pressure, mean (SD)	Systolic	108 (17)	102 (16)
	Diastolic	69 (13)	66 (11)
Pulse, mean (SD), beats/minute		73 (11)	73 (13)
Blood test values, mean (SD)	Creatinine, umol/L	108 (34)	105 (41)
	Sodium, mmol/L	139 (3)	139 (3)
	Potassium, mmol/L	4.4 (0.4)	4.2 (0.5)
	Hemoglobin, g/L	135 (13)	142 (15)
	Urate, umol/L	413 (117)	412 (124)
Duration of heart failure, median, (interquartile range), years		4.8 (7.8)	3.5 (8.2)
Primary cause of heart failure, No. (%)	Ischemic	20 (40)	13 (26)

Characteristic		telemonitoring group (N = 50)	standard care group (N = 50)
	Idiopathic	22 (44)	29 (58)
	Other	8 (16)	8 (16)

[a]Missing values account for totals less than 100%.

Figure 3. Weekly adherence to completing all daily measurements.

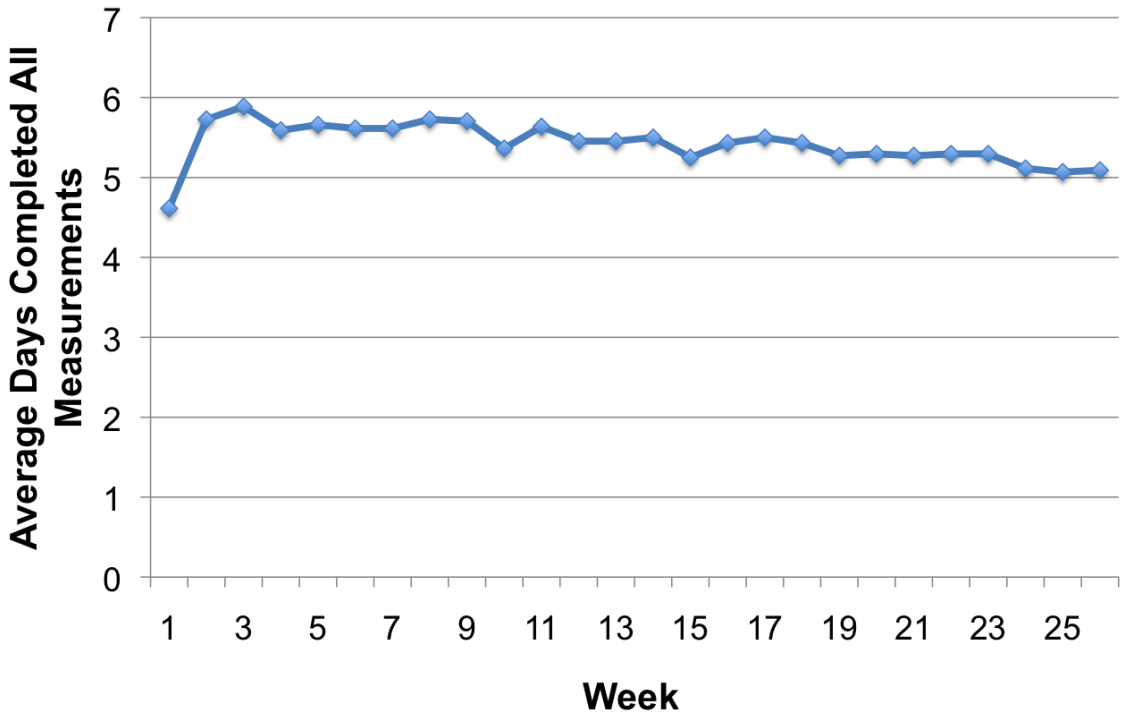


Table 2. Clinical utilization of the telemonitoring system and clinical actions resulting from alerts

Clinical system use/action	Times occurred over the 6-month trial
Email alerts sent to Cardiologist A; Cardiologist B	1367; 311
Logged onto website by Cardiologist A; Nurse Practitioner A	67; 34
Phoned patient due to alerts	480
Medication changed or medication instructions given	105
Ordered additional blood work	26
Moved clinic visit forward to an earlier date	9
Instructed patient to go to local emergency department	6
Instructed patient to contact family physician	4

BNP, NYHA Class, LVEF, SCHFI, and MLHFQ Results

Table 3 shows the results from the statistical analysis. BNP values for both the telemonitoring and standard care groups decreased post-study (P = .001, P = .002, respectively). Similarly, NYHA class, LVEF, self-care maintenance, and self-care management improved for both groups. However, quality of life as measured with the MLHFQ significantly improved only for the telemonitoring group (P = .02), including the physical (P = .02) and emotional dimensions (P = .03).

A comparison of the post-study data between groups found only a statistically significant difference in SCHFI maintenance scores, indicating the telemonitoring group had greater self-care maintenance (ie, a higher SCHFI maintenance score) (P = .03). A comparison of the change scores between groups found only a statistically significant difference in the overall MLHFQ scores, indicating that the telemonitoring group had greater improvement in quality of life (ie, a larger difference from baseline to post-study MLHFQ scores) (P = .05).

Table 3. Results for BNP, NYHA class, LVEF, SCHFI, and MLHFQ scores

Parameter	telemonitoring group				standard care group				Between-group post-study data P value	Between-group change scores P value
	N	Baseline mean (SD)	Post-study mean (SD)	P value	N	Baseline mean (SD)	Post-study mean (SD)	P value		
BNP (pg/mL)	44	592 (538)	414 (604)	.001	44	426 (501)	303 (460)	.002	.2	.5
NYHA class	43	2.5 (0.6)	2.1 (0.7)	.000	38	2.6 (0.6)	2.2 (0.7)	.001	.8	.8
LVEF (%)	41	25.2 (8.8)	32.7 (11.8)	.001	35	24.8 (9.7)	31.3 (12.5)	.001	.7	.7
Self-care maintenance (SCHFI)	38	65.1 (18.6)	73.3 (11.6)	.004	44	58.9 (18.7)	65.5 (15.8)	.006	.03	.6
Self-care management (SCHFI)	18	58.1 (24.5)	68.6 (16.0)	.02	21	57.9 (22.4)	69.3 (18.3)	.01	.7	.9
Self-care confidence (SCHFI)	37	57.4 (20.6)	57.7 (19.5)	.7	43	55.8 (20.0)	56.2 (21.8)	.9	.9	.8
Quality of life (MLHFQ)	38	50.3 (29.1)	41.4 (26.7)	.02	44	47.8 (22.6)	47.3 (23.4)	.9	.2	.05
Quality of life – physical (MLHFQ)	38	21.7 (12.8)	17.8 (12.9)	.02	44	21.0 (10.5)	20.2 (10.5)	.5	.3	.1
Quality of life – emotional (MLHFQ)	38	11.7 (8.6)	9.5 (7.8)	.03	44	11.0 (7.0)	11.3 (6.9)	.8	.2	.07

As with previous studies of multidisciplinary heart function clinics, it was hypothesized being enrolled in the Heart Function Clinic would improve outcomes particularly for those patients new to the clinic. To test if there was a clinic effect on the outcome measures showing improvements in both groups, the changes in outcome measures were compared for patients who were new to the clinic (< 6 months, N = 37) versus those who were long-term patients (> 6 months, N = 63). We chose six months as the cut-off point because patients newly referred to the clinic usually require three months to up-titrate medication and a further three months to reach clinical stability. In the telemonitoring group, 18 patients were new to the clinic versus 19 in the standard care group. Patients who were new to the clinic improved more than those who were long-term with respect to BNP and LVEF (Table 4). Self-care maintenance of the new patients also improved more, but the difference was not statistically significant.

Table 4. Comparison of patients new to the Heart Function Clinic (enrolled < 6 months) and long-term patients (enrolled > 6 months)

Parameter	Patients enrolled < 6 months		Patients enrolled > 6 months		P value
	N	Baseline mean (SD)	N	Baseline mean (SD)	
Change in BNP (pg/mL)	35	279.8 (606.4)	53	65.4 (325.5)	.003
Change in NYHA class	33	0.5 (0.8)	44	0.4 (0.5)	.6
Change in LVEF (%)	34	−9.5 (10.6)	42	−5.1 (11.4)	.02
Change in self-care maintenance (SCHFI)	31	−10.6 (17.9)	51	−5.4 (14.1)	.1
Change in self-care management (SCHFI)	12	−9.5 (17.1)	27	−11.7 (19.4)	.7

To minimize the clinic effect, the statistical analysis was repeated post-hoc removing the data from the 37 patients new to clinic. Table 5 shows the subgroup statistical results for the parameters that improved for both groups. For long-term patients, the BNP (P = .02), LVEF (P = .005), self-care maintenance (P = .05), and self-care management (P = .03) significantly improved only for the telemonitoring group. A comparison of the post-study data and change scores between groups found only a statistically significant difference in post-study MLHFQ emotional dimension scores, indicating the telemonitoring group had better post-study quality of life (emotional dimension) compared to the standard care group (P = .05).

Table 5. Results for BNP, NYHA class, LVEF, and SCHFI scores with new patients to the clinic removed

Parameter	telemonitoring group				standard care group				Between-group post-study data *P* value	Between-group change scores *P* value
	N	Baseline mean (SD)	Post-study mean (SD)	*P* value	N	Baseline mean (SD)	Post-study mean (SD)	*P* value		
BNP (pg/mL)	27	583 (464)	433 (445)	.02	26	326 (296)	349 (467)	.4	.3	.1
NYHA Class	26	2.6 (0.5)	2.2 (0.7)	.002	18	2.7 (0.6)	2.3 (0.5)	.01	1.0	.9
LVEF (%)	24	23.9 (8.9)	31.3 (12.3)	.005	18	28.1 (10.5)	30.1 (14.7)	.4	.9	.1
Self-care maintenance (SCHFI)	23	64.9 (20.3)	71.9 (12.7)	.05	28	59.1 (18.7)	63.1 (16.9)	.1	.07	.5
Self-care management (SCHFI)	13	51.5 (23.5)	65.0 (12.9)	.03	14	62.1 (17.8)	72.1 (12.0)	.08	.2	.7

Mortality Results

During the trial, three patients from the telemonitoring group died and none died from the standard care group. One patient died from newly diagnosed cancer, the second was suspected to have died from sepsis due to a leg ulcer, and the third died from post-heart transplantation complications (transplant performed prior to study enrollment).

Health Care Resource Utilization Results

Table 6 shows the results from the Mann–Whitney tests comparing the health care resource utilization by the telemonitoring and standard care groups during the trial. No differences were found between the groups for number of hospitalizations ($P = .1$), number of nights in hospital ($P = .2$), or number of visits to the emergency room ($P = .6$). However, the telemonitoring group visited the Heart Function Clinic during the six months more often than the standard care group ($P = .04$) because the cardiologist managing the alerts asked several patients to come into clinic when their health appeared to be deteriorating, as alerted by the telemonitoring system.

Table 6. Results for health care resource utilization

Parameter	telemonitoring group		standard care group		*P* value
	N	Mean (SD)	N	Mean (SD)	
Number of hospital admissions	38	0.5 (0.8)	44	0.2 (0.4)	.1
Number of nights in hospital	38	2.3 (5.3)	44	1.3 (4.2)	.2
Number of emergency department visits	38	0.4 (0.9)	44	0.3 (0.5)	.6
Number of Heart Function Clinic visits	39	3.5 (3.6)	45	2.5 (2.5)	.04

Medication and Blood Test Results

Table 7 shows the number of patients prescribed angiotensin-converting enzyme (ACE) inhibitor and/or angiotensin receptor blocker (ARB), beta-blocker, diuretic, statin, aldosterone antagonist, digoxin, and antiarrhythmic medication at baseline and post-study for the telemonitoring and standard care groups. Using McNemar's test, no statistical differences were found between the frequency of medication prescribed at baseline compared to the frequency of medication prescribed post-study for either of the groups, except more of the patients in the telemonitoring group were prescribed aldosterone antagonist post-study ($P = .02$). Seven patients in the telemonitoring group started to take aldosterone antagonist during the six-month trial. No significant statistical differences were found comparing the post-study frequency of medication between groups using Fisher's exact test. No significant differences were found between baseline and post-study creatinine, sodium, potassium, hemoglobin, and urate blood test results.

Table 7. Number of patients prescribed various types of medication

Medication	telemonitoring group					standard care group					Between-group post-study data *P* value
	Baseline		Post study		*P* value	Baseline		Post study		*P* value	
	Pre-scribed	Not pre-scribed	Pre-scribed	Not pre-scribed		Pre-scribed	Not pre-scribed	Pre-scribed	Not pre-scribed		
ACE-inhibitor and/or ARB	49	1	44	1	1.0	48	2	36	2	1.0	.6
Beta-blocker	49	1	44	1	1.0	49	1	38	0	1.0	1.0
Diuretic	47	3	42	3	1.0	45	5	34	4	1.0	.7
Statin	31	19	27	18	1.0	26	24	22	16	.3	1.0
Aldosterone antagonist	23	27	27	18	.02	29	21	23	15	.7	1.0
Digoxin	16	34	16	29	1.0	22	28	16	22	.7	.7
Antiarrhythmic	14	36	12	33	1.0	6	44	6	32	1.0	.3

Discussion

A randomized controlled trial was performed to evaluate a user-centered, mobile phone-based telemonitoring system. Although the trial was underpowered to detect its impact on hospitalization and mortality, the results suggest quality of life improved with the use of the system through increased self-care and improved clinical management. BNP, LVEF, and NYHA class all improved over the course of the trial for both the telemonitoring and standard care groups. A subgroup analysis using only the participants who had attended the clinic for more than 6 months showed only the telemonitoring group had significant improvements in BNP and LVEF from baseline to post-study. It is possible a trial with a larger sample size would find a reduction in hospitalization and mortality in the telemonitoring group.

One of the most significant changes in clinical management with the telemonitoring system was the ability to optimize a patient's medication regimen. For example, there was a statistically significant increase in the number of patients in the telemonitoring group who were prescribed aldosterone antagonist compared to the standard care group. Previous studies have found less than a third of eligible patients receive heart failure guideline-recommended aldosterone antagonist therapy [15]. The benefits of aldosterone antagonist, in terms of reductions in mortality and hospitalizations, have been well documented. However, patients are often not prescribed this therapy partially because of the need to closely monitor serum potassium levels due to the risk of hyperkalemia [16]. It is possible the increase in use of aldosterone antagonist in the telemonitoring group can be attributed to the close monitoring of the patients enabled by the telemonitoring system.

In terms of self-care, the telemonitoring system provided immediate automated instructions and enabled clinical intervention at the most appropriate time ("teachable moments") to help patients modify their lifestyle behaviors. For example, many patients found their weight and blood pressure increased after a high sodium meal. By reducing their salt intake for the next few meals, their weight and blood pressure would return to within their normal range. Some patients also received automated reminders to take extra diuretic medication after a weight gain, which confirmed taking the extra medication was the correct course of action. Even though this group of patients received prior instruction from their cardiologist to take extra diuretic medication in this situation, they were still often hesitant to take the extra medication without prior confirmation.

A large-scale (N=1653) randomized controlled trial, Telemonitoring to Improve Heart Failure Outcomes (Tele-HF), was conducted using as a primary outcome measure a composite of readmission for any reason or death from any cause within 180 days of enrollment [17]. That trial found no reduction in mortality or hospital admissions from the use of a telephone-based interactive voice-response system to record heart failure symptoms and weight data. One of the possible reasons no differences were found between groups is the Tele-HF trial attempted to engage patients in self-care only in terms of performing the daily reporting of symptoms and weight. Real-time automated self-care advice and instructions (taking extra diuretic medication, following salt and fluid restrictions, etc) based on the reported symptoms and weight—as implemented in our study—might have had a significant positive impact. In addition, site coordinators reviewed the patient information daily on weekdays for the Tele-HF trial and should have contacted the patient as required. It is possible the necessary real-time response by clinicians was not provided.

Finally, 14% of the Tele-HF patients randomized into the intervention group never used the system, and only 55% of the patients were using the system at least 3 times per week by the final week. The Tele-HF trial lasted 180 days, which was similar to our trial. However, by the final week of our trial, 89% of our patients were taking their measurements at least 3 times per week (excluding the 3 patients who had died). Our high rate of adherence may be due to the perceived benefit, ongoing positive reinforcement, and ease of use of our system even among the very elderly (the oldest patient in the telemonitoring group was aged 88 years). The trial by Scherr et al (2009) also found low adherence rates due to patients having difficulty in manually entering and sending their daily blood pressure, heart rate, body weight, and dosage of heart failure medication through a mobile phone's Internet browser [6]. The differences in adherence rates

between these trials underscore the importance of system design on the successful implementation of telemonitoring systems.

The TIM-HF trial investigated the effects of wireless medical devices with a personal digital assistant (PDA) on health outcomes of 710 heart failure patients randomized into intervention and control groups. Similar to the Tele-HF trial, this trial found no statistically significant differences between the groups with respect to mortality, number of hospitalizations, or days in hospital [5]. However, an exploratory subgroup analysis suggested the efficacy of telemonitoring to improve heart failure outcomes could be dependent on the characteristics of the patient population. The investigators concluded further trials were required to explore the effects of telemonitoring in defined patient subgroups.

In our study, being newly enrolled into the Heart Function Clinic overshadowed improvements from the telemonitoring system in terms of the within-group analysis. In addition, many of our participants had stable heart failure, and had not been admitted to hospital for many years. Bowles et al (2009) also suggested targeting higher risk patients with greater probability of being rehospitalized in order to demonstrate the most effective outcomes from telemonitoring [18]. Eligibility criteria for study participation that included admission to hospital within the previous year and being part of the Heart Function Clinic for at least 6 months may have resulted in more significant improvements in health outcomes.

Although the clinicians viewed the ECG recordings that were available, the benefits of the ECG recordings compared to the costs of the devices were not conclusive. Feedback from the clinicians indicated they thought the ECG recordings were of some use, but the inability to provide the devices to participants with ICDs was a significant drawback.

Limitations

A confounder to our study was the clinic effect that caused improvements in outcomes in both the telemonitoring and standard care groups, as described above. Future studies should consider recruitment of stable clinic patients. Secondly, patients were enrolled in the winter and completed the trial in the summer when heart failure patients are often healthier due to reductions in respiratory illnesses. This seasonal effect may have also contributed to the observed improvements in both the telemonitoring and standard care groups. A third limitation was the small sample size that did not provide adequate power to detect the effects of the telemonitoring on mortality and hospitalization outcomes. A future trial with a larger sample size would help further determine the effects of the telemonitoring system on health service utilization and its health economic impact. In addition, the proportionate benefit of self-care versus clinical management changes on health outcomes could not be definitively determined. For example, the improvements observed in the telemonitoring group may be largely attributed to the increased prescription of aldosterone antagonist. Furthermore, about a third of the patients in the telemonitoring group used the telemonitoring system for a number of weeks prior to completing the baseline questionnaire. Although the patients were instructed to answer the questionnaire based on information before they used the telemonitoring system, it was clear many patients were basing their answers on information associated with system use because their questionnaire responses sometimes did not match the information provided during their recruitment interviews. This minimized the measured impact of the telemonitoring system. Limitations to the questionnaire data also include potential recall bias and self-reporting. Finally, the telemonitoring system was only available in English, which limited the participant population to patients who were able to read rudimentary English.

Conclusions

The results from our trial suggest mobile phone-based telemonitoring improves quality of life through improved self-care and clinical management. A subgroup analysis using only the participants who had attended clinic for more than 6 months showed only the telemonitoring group had significant improvements in BNP and LVEF from baseline to post-study. An important component to successful telemonitoring for heart failure appears to be immediate feedback to the patients to address any potential decompensation either through automated messages and/or advice from a clinician who is familiar with patients' histories. In addition, in order for patients to be willing to integrate telemonitoring into their daily lives, the system must be easy and quick to use. Further research with an appropriate heart failure patient population and larger sample size is required to determine the extent of the benefits of such a telemonitoring system on heart failure outcomes.

Acknowledgments

Funding for this work was in part provided by the Toronto General Hospital Foundation and the Natural Sciences and Engineering Research Council of Canada Strategic Research Network Grant entitled Healthcare Support through Information Technology Enhancements (hSITE).

Conflicts of Interest

None

References

1. Inglis SC, Clark RA, McAlister FA, Ball J, Lewinter C, Cullington D, et al. Structured telephone support or telemonitoring programmes for patients with chronic heart failure. Cochrane Database Syst Rev 2010(8):CD007228. [doi: 10.1002/14651858.CD007228.pub2] [Medline: 20687083]

2. Clark RA, Inglis SC, McAlister FA, Cleland JG, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. BMJ 2007 May 5;334(7600):942 [FREE Full text] [doi: 10.1136/bmj.39156.536968.55] [Medline: 17426062]
3. Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K, et al. Home telemonitoring for congestive heart failure: a systematic review and meta-analysis. J Telemed Telecare 2010;16(2):68-76. [doi: 10.1258/jtt.2009.090406] [Medline: 20008054]
4. Klersy C, De Silvestri A, Gabutti G, Regoli F, Auricchio A. A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 2009 Oct 27;54(18):1683-1694. [doi: 10.1016/j.jacc.2009.08.017] [Medline: 19850208]
5. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, Telemedical Interventional Monitoring in Heart Failure Investigators. Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure: the telemedical interventional monitoring in heart failure study. Circulation 2011 May 3;123(17):1873-1880. [doi: 10.1161/CIRCULATIONAHA.111.018473] [Medline: 21444883]
6. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, MOBITEL Investigators. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009;11(3):e34 [FREE Full text] [doi: 10.2196/jmir.1252] [Medline: 19687005]
7. Riegel B, Vaughan Dickson V, Goldberg LR, Deatrick JA. Factors associated with the development of expertise in heart failure self-care. Nurs Res 2007 Jul;56(4):235-243. [doi: 10.1097/01.NNR.0000280615.75447.f7] [Medline: 17625462]
8. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the self-care of heart failure index. J Cardiovasc Nurs 2009 Dec;24(6):485-497. [doi: 10.1097/JCN.0b013e3181b4baa0] [Medline: 19786884]
9. Urbaniak GC, Plous S. Research Randomizer. URL: http://www.randomizer.org/ [accessed 2011-11-16] [WebCite Cache ID 63FKn7qtR]
10. Riegel B, Moser DK, Glaser D, Carlson B, Deaton C, Armola R, et al. The Minnesota Living With Heart Failure Questionnaire: sensitivity to differences and responsiveness to intervention intensity in a clinical population. Nurs Res 2002 Aug;51(4):209-218. [Medline: 12131233]
11. Riegel B, Carlson B, Moser DK, Sebern M, Hicks FD, Roland V. Psychometric testing of the self-care of heart failure index. J Card Fail 2004 Aug;10(4):350-360. [Medline: 15309704]
12. Rector TS, Cohn JN. Minnesota Living With Heart Failure Questionnaire. URL: http://www.license.umn.edu/Products/Minnesota-Living-With-Heart-Failure-Questionnaire__Z94019.aspx [accessed 2011-08-08] [WebCite Cache ID 60nHNOGV7]
13. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
14. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and Experiences of Heart Failure Patients and Clinicians on the Use of Mobile Phone-Based Telemonitoring. J Med Internet Res 2012 Oct 21;14(1):e25-e65 [FREE Full text] [doi: 10.2196/jmir.1912]
15. Albert NM, Yancy CW, Liang L, Zhao X, Hernandez AF, Peterson ED, et al. Use of aldosterone antagonists in heart failure. JAMA 2009 Oct 21;302(15):1658-1665. [doi: 10.1001/jama.2009.1493] [Medline: 19843900]
16. Zannad F, McMurray JJ, Krum H, van Veldhuisen DJ, Swedberg K, Shi H, EMPHASIS-HF Study Group. Eplerenone in patients with systolic heart failure and mild symptoms. N Engl J Med 2011 Jan 6;364(1):11-21 [FREE Full text] [doi: 10.1056/NEJMoa1009492] [Medline: 21073363]
17. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
18. Bowles KH, Holland DE, Horowitz DA. A comparison of in-person home care, home care with telephone contact and home care with telemonitoring for disease management. J Telemed Telecare 2009;15(7):344-350. [doi: 10.1258/jtt.2009.090118] [Medline: 19815903]

Abbreviations

ACE: angiotensin-converting enzyme
ARB: angiotensin receptor blocker
BNP: brain natriuretic peptide
ECG: electrocardiograms
ICD: implantable cardioverter defibrillator
LVEF: left ventricular ejection fraction
MLHFQ: Minnesota Living with Heart Failure Questionnaire
NYHA: New York Heart Association
RCT: randomized controlled trial
SC: standard care
SCHFI: Self-Care of Heart Failure Index

SD: standard deviation
Tele-HF: Telemonitoring to Improve Heart Failure Outcomes
TIM-HF: Telemedical Interventional Monitoring in Heart Failure
telemonitoring: telemonitoring
UHN: University Health Network

Edited by H Potts; submitted 09.08.11; peer-reviewed by J Polisena, A Kollmann; comments to author 19.09.11; revised version received 04.12.11; accepted 12.12.11; published 16.02.12.

Please cite as:
Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ
Mobile Phone-Based Telemonitoring for Heart Failure Management: A Randomized Controlled Trial
J Med Internet Res 2012;14(1):e31
URL: http://www.jmir.org/2012/1/e31/
doi:10.2196/jmir.1909
PMID:22356799

Viewpoint

Developing Health Promotion Interventions on Social Networking Sites: Recommendations from The FaceSpace Project

Judy Gold[1,2], PhD; Alisa E Pedrana[1,2], BBiomedSci(Hons); Mark A Stoove[1,2], PhD; Shanton Chang[3], PhD; Steve Howard[3], PhD; Jason Asselin[4], GradCert(HlthProm); Olivia Ilic[5], GradDip(Animateuring); Colin Batrouney[4], BA; Margaret E Hellard[1,2,6], PhD

[1]Centre for Population Health, Burnet Institute, Melbourne, Australia

[2]Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia

[3]Department of Information Systems, The University of Melbourne, Melbourne, Australia

[4]Victorian AIDS Council/Gay Men's Health Centre, Melbourne, Australia

[5]X:MACHINE Productions, Melbourne, Australia

[6]The Nossal Institute for Global Health, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Judy Gold, PhD
Centre for Population Health
Burnet Institute
85 Commercial Rd
Melbourne, 3004
Australia
Phone: 61 3 8506 2324
Fax: 61 3 9282 2138
Email: judy@burnet.edu.au

Abstract

Online social networking sites offer a novel setting for the delivery of health promotion interventions due to their potential to reach a large population and the possibility for two-way engagement. However, few have attempted to host interventions on these sites, or to use the range of interactive functions available to enhance the delivery of health-related messages. This paper presents lessons learnt from "The FaceSpace Project", a sexual health promotion intervention using social networking sites targeting two key at-risk groups. Based on our experience, we make recommendations for developing and implementing health promotion interventions on these sites. Elements crucial for developing interventions include establishing a multidisciplinary team, allowing adequate time for obtaining approvals, securing sufficient resources for building and maintaining an online presence, and developing an integrated process and impact evaluation framework. With two-way interaction an important and novel feature of health promotion interventions in this medium, we also present strategies trialled to generate interest and engagement in our intervention. Social networking sites are now an established part of the online environment; our experience in developing and implementing a health promotion intervention using this medium are of direct relevance and utility for all health organizations creating a presence in this new environment.

(J Med Internet Res 2012;14(1):e30) doi:10.2196/jmir.1875

KEYWORDS

Health promotion; Internet; social networking sites

Introduction

Over the past 20 years the Internet has dramatically changed how individuals access information and communicate. Global Internet use has grown exponentially, with an estimated 1.8 billion Internet users in 2009, up from 318 million users in 1998 [1]. The Internet is increasingly used for health purposes [2]; one survey reported 83% of American Internet users source health information online [3]. Numerous Internet-based health interventions have been developed, with several reviews concluding that such interventions generally have positive effects for a range of behaviours [4-7].

'Web 2.0' is a relatively recent development that refers to a loose collection of web-based technologies and services that

allow end users to interact and collaborate as content creators, rather than the one-way information flow on relatively static 'Web 1.0' websites [8-10]. The term 'social media' is used interchangeably with Web 2.0 to describe sites and applications that allow information sharing and interactive activities among online communities; examples include blogs, wiki's, content-sharing sites, virtual worlds and social networking sites [10,11].

Social networking sites allow individuals to maintain, form and visualize their social networks, and often offer additional functions such as public and private messaging and photo, video and other content sharing [12]. Facebook, Twitter, LinkedIn and MySpace are the most popular social networking sites globally [13], with others largely popular only within certain sub-groups or geographical regions [12]. Growth in usage has been extremely rapid, with Facebook reporting 500 million active users [14], up from 200 million in April 2009 [15].

Commercial organizations have been quick to capitalize on the utility of using Web 2.0 to attract, retain and engage end users [10], while health organizations have lagged behind [10,16,17]. Very little has been published about how social networking sites might be exploited for health promotion interventions. A recent review of the use of social media for social marketing identified just four examples, none of which used the most common social networking sites listed earlier [18]. Some health organizations have begun extending their presence into social networking sites [19-22]; however, this has often been used as an additional form of marketing to promote services rather than for intervention delivery. Other work has focused on the public display of risky behaviour (e.g. alcohol use) on these sites [23,24]. However, there are few published examples of organizations actually *delivering* health promotion interventions through social networking sites.

The lack of published examples describing intervention delivery using social networking sites makes it very difficult for others to realistically consider if and how they might approach developing interventions in these spaces. Moreover, the lack of evidence for evaluating such interventions makes it difficult to determine if health promotion interventions using social networking sites are effective.

During 2009 and 2010, we implemented a novel health promotion intervention using social networking sites: "The FaceSpace Project". The aim of this paper is to use our experience to provide recommendations for developing health promotion interventions on social networking sites.

The FaceSpace Project

The FaceSpace Project trialled the delivery of sexual health promotion via social networking sites to two key at-risk groups: young people aged 16-29 years, and men who have sex with men (MSM). The project concept was to use fictional characters to post content (primarily videos) and to interact on various social networking sites, with sexual health promotion messages embedded within some postings and interactions. The project was a collaboration between public health researchers, experts in user interaction with information technologies, a creative productions company, and a community organization.

The young people's arm was developed and implemented first. Two young male and two young female characters and character narratives were developed in workshops with young people, actors, and project staff, and character narratives developed. Each character had a Facebook page (www.facebook.com/thefacespaceproject), and a presence on one other social networking site (Twitter, Flickr, YouTube) (www.youtube.com/thefacespaceproject) (Figure 1). (Note that the pages for the young people's arm are no longer actively maintained.) The overall project also had a Facebook page and a YouTube channel. From November 2009 until April 2010, each character posted regular updates and periodic videos on their sites, including interactions on each other's sites. Project evaluation included site usage and interaction statistics, questionnaires and focus group discussions.

The learnings from the young people's arm of the project informed the development of the MSM arm. This arm was separately branded 'Queer As F**k' and launched online in April 2010. In this arm, four male characters (all MSM) were developed; however, the emphasis of the development phase was on the series narrative (rather than character-focused), and all the characters interacted together on one Facebook page (www.facebook.com/QAFxxk), supported by one YouTube channel (www.youtube.com/queerasfxxk) (Figure 2). Unlike the youth arm where the videos were styled predominantly as personal blogs, videos in the MSM arm were episodic in nature with a cohesive narrative and sexual health themes embedded in most episodes. Similar evaluation methods were used to the youth arm, with the addition of a diary-scrapbook.

Figure 1. Screenshots from the arm of The FaceSpace Project targeting young people.

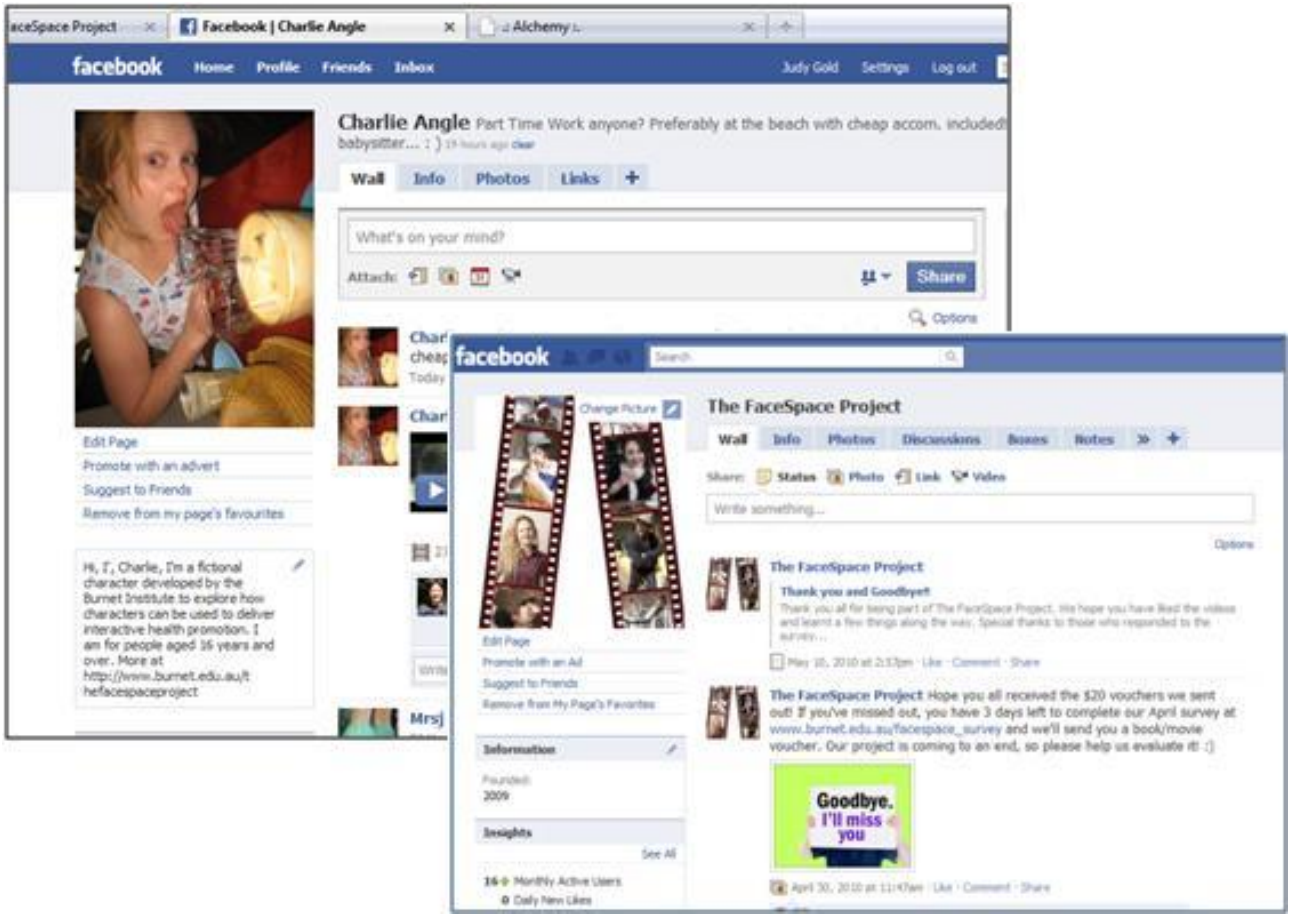
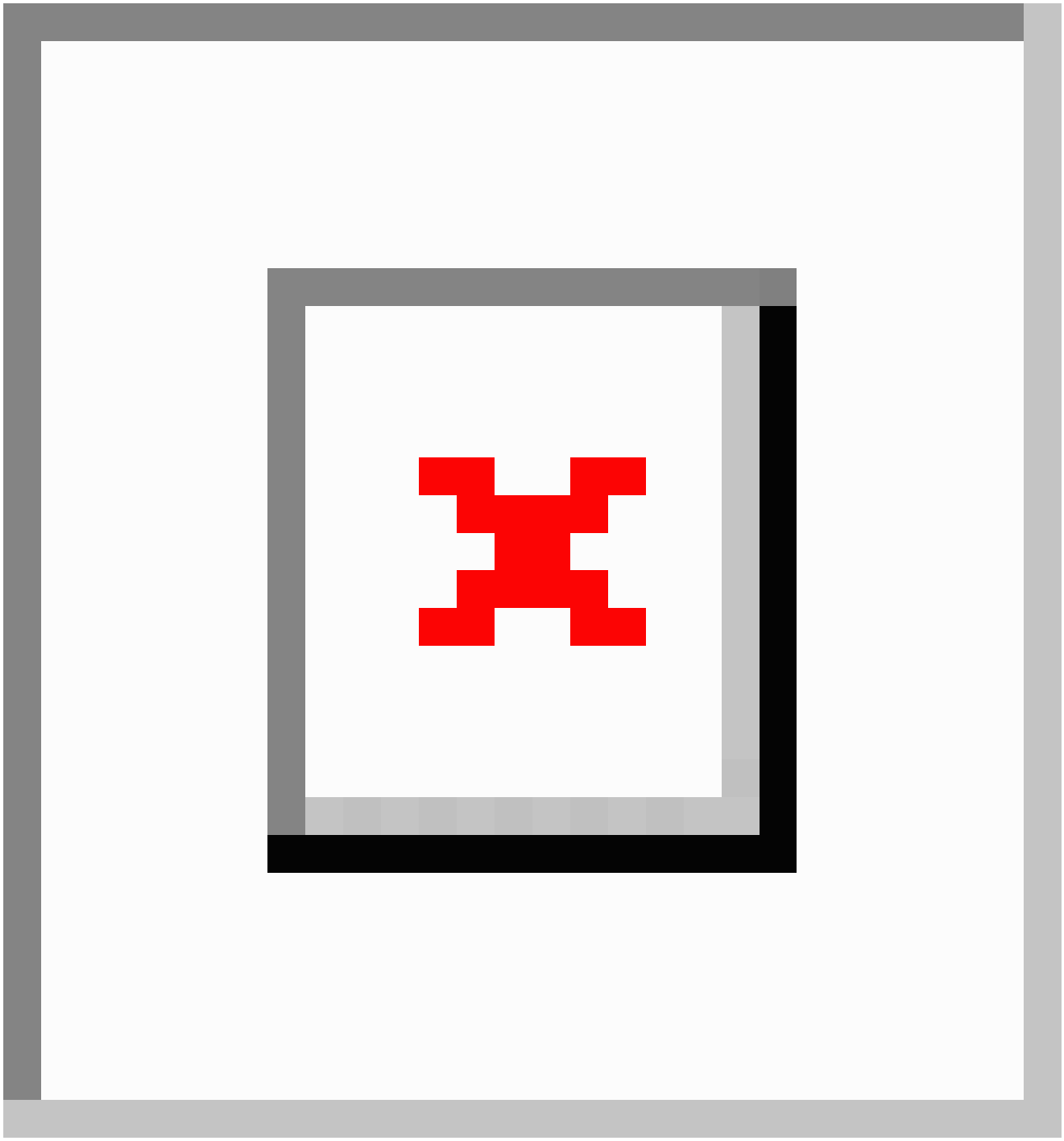


Figure 2. Screenshots from the arm of The FaceSpace Project (Queer As F**k) targeting men who have sex with men.



Key Outcomes From The FaceSpace Project

At the conclusion of the young people's arm of the project, the 5 Facebook pages had a total of 900 fans. The 31 project videos had 5300 total views on YouTube, with views of individual videos ranging from 12 to 3188 views. Interaction on the Facebook pages varied over time (Figure 3), with peaks generally corresponding to posting of project videos.

At the conclusion of the arm of the project targeting gay men ('Queer as F**k'), the Facebook page had 1332 fans. The 10 video episodes of the project had 7886 views, with views of individual episodes ranging from 256 to 1814 views. As with the youth arm of the project, interaction on the Facebook page varied over time, with peaks in interactions generally corresponding to when new video episodes were posted (Figure 4).

Since the conclusion of The FaceSpace Project, the arm targeting gay men ('Queer as F**k') has been taken up by the project's key community partner, the Victorian AIDS Council/Gay Men's Health Service. Subsequent seasons of Queer as F**k now form an integral part of their social marketing campaigns. Findings from The FaceSpace Project have guided modifications to these subsequent seasons by using more implicit sexual health messages embedded within dramatic threads, and by creating an online environment for organic user-led, rather than expert-led, dialogue.

Figure 3. Daily number of interactions ('likes', wall posts and comments) on the Facebook pages of the arm of The FaceSpace Project targeting young people.

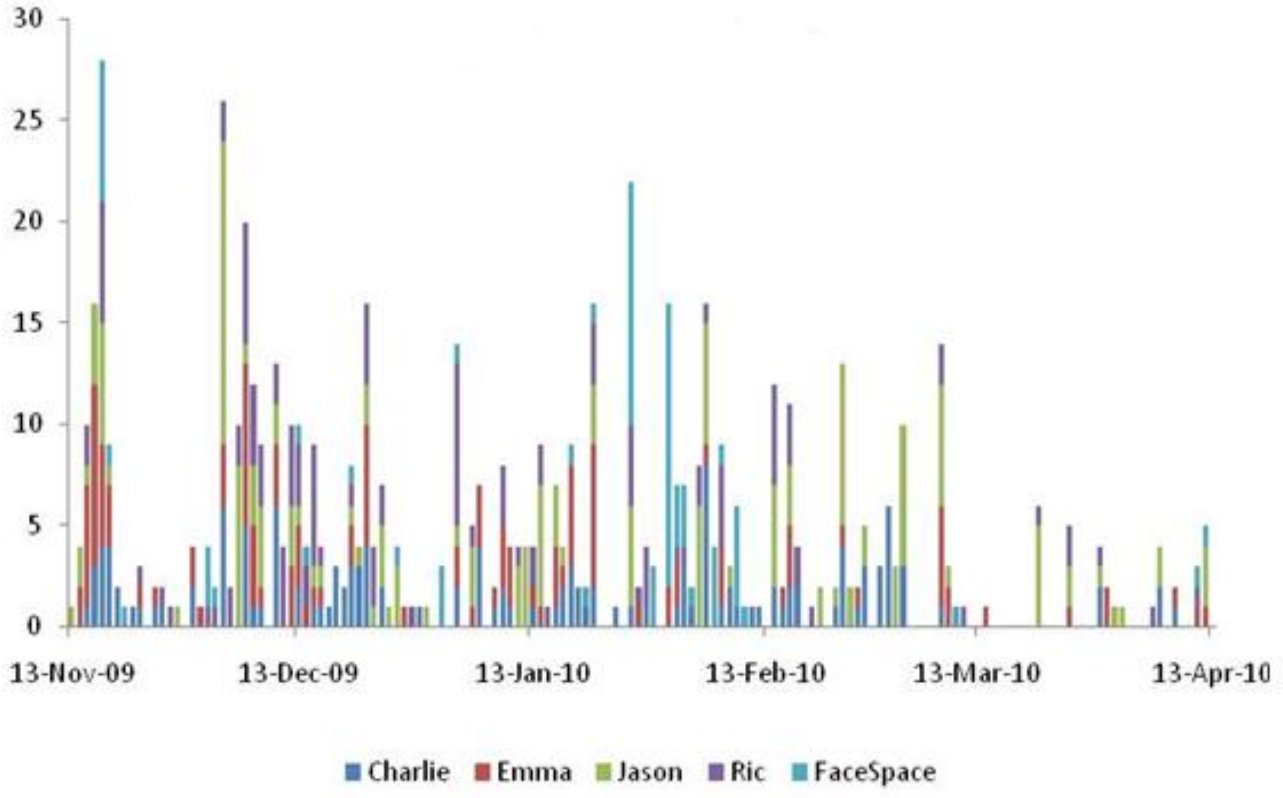
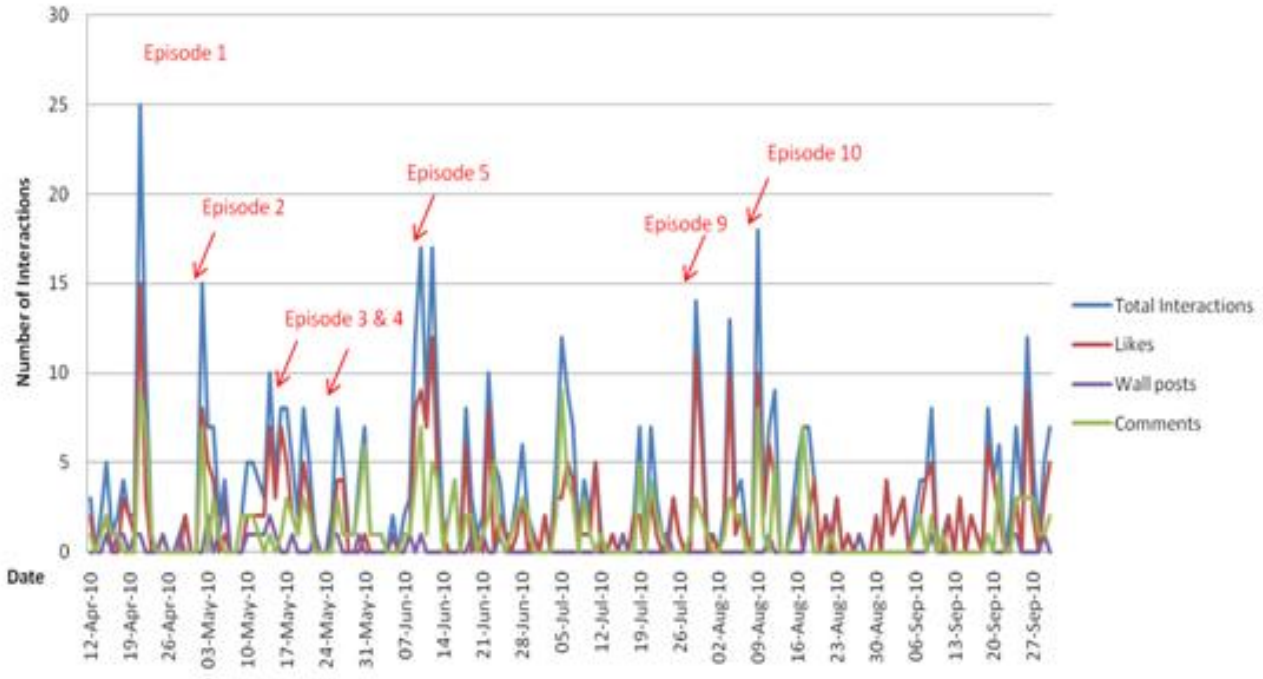


Figure 4. Daily number of interactions ('likes', wall posts and comments) on the Facebook page of the arm of The FaceSpace Project targeting men who have sex with men (Queer As F**k).



Key Recommendations from The FaceSpace Project

Key recommendations from the FaceSpace Project are shown in Table 1.

Table 1. Recommendations from The FaceSpace Project.

	Recommendation
1	Create and nurture a multidisciplinary team with all the skills required—just because you can drive a car and change the oil doesn't make you a mechanic
2	Anticipate delays getting approval (ethical, legal, organizational)—it's a new medium and sometimes the waters haven't been tested
3	Resource, resource, resource—you will need time, money, human and brain power to develop and maintain sites (without forgetting the rest)
4	Generate interest (buzz) and do it early—just because you've built it, doesn't mean they'll come
5	Keep your audience engaged—don't fall off the newsfeed!
6	Go viral—if you find the formula, you're a millionaire
7	Define success and how you will measure it

Recommendation 1. Create and Nurture a Multidisciplinary Team With All the Skills Required—Just Because You Can Drive a Car and Change the Oil Doesn't Make You a Mechanic.

Unlike standard health promotion interventions where many organizations have the in-house expertise required for implementation, interventions on social networking sites require additional expertise in social media and knowledge of how end users interact and engage in online environments. Familiarity with social networking sites from personal experience is insufficient to build and maintain an organization presence or to design a health promotion intervention in these spaces. Teams require a broad range of skills and knowledge, including an adequate understanding of potential sites and their functionality as well as an understanding of social marketing.

We formed a multidisciplinary project team that involved public health researchers (Burnet Institute), experts in how end users interact with technology (Department of Information Systems, University of Melbourne), a creative productions company experienced in online performances (X:MACHINE) and a community organization experienced in sexual health promotion (Victorian AIDS Council/Gay Men's Health Centre). An advisory group comprised of experts in various fields related to the project was also convened to provide ongoing advice.

Although our multidisciplinary project team was successfully established, such collaborations bring difficulties of their own, including ensuring timely and adequate communications, clear delineation of roles and responsibilities, and interdisciplinary tensions including different philosophies underpinning approaches to design and implementation (eg, user-led vs creative-led design). This was the first time this team had worked together, and we had not anticipated the resources (time, financial) required to build and maintain this collaboration. Such resourcing is vital to ensure a healthy and vibrant collaboration to support the development of effective interventions.

Recommendation 2. Anticipate Delays Getting Approval (Ethical, Legal, Organizational)—It's a New Medium and Sometimes the Waters Haven't Been Tested.

The use of social networking sites for health promotion interventions can raise ethical, legal and organizational concerns. In addition, individuals and boards who are responsible for approving interventions may not be familiar with social networking sites or how they are used by individuals [25]. Potential concerns include privacy, consent, intervention access, duty of care, organizational reputation, data collection and management, and reduced control over message delivery compared to other settings.

In our case, while legal approval was relatively straightforward, we had some challenges negotiating intellectual property ownership between the collaborating organizations. In addition, we underwent a lengthy review process before being granted approval by our ethical review board. One positive outcome of this review included development of a clearer and more detailed protocol for responding to 'inappropriate' posts on our pages (see Multimedia Appendix 1). However, we were required to significantly modify the delivery of our intervention in several ways. For example, the board required prominent disclaimers on the page and regular reminders to fans that reinforced that our characters were fictional, and warnings to not post information that individuals 'may regret later'. We believe these requirements may have negatively impacted on our credibility on social networking sites and thus reduced end users' willingness to participate and engage with our intervention.

Social networking sites are a new and challenging environment for many organizations. This should be anticipated in project timelines, as applying and obtaining ethical, legal and organizational approval can be time-consuming and difficult. Content areas considered socially 'sensitive' (such as ours) or related to illicit behaviour (eg, drug use) may attract additional scrutiny, given the public nature of social networking sites. Including a "Social Networking 101" education component for approval bodies during the development period may be a useful strategy to minimise delays in obtaining approval.

Recommendation 3. Resource, Resource, Resource—You Will Need Time, Money, Human and Brain Power to Develop and Maintain Sites (Without Forgetting the Rest).

One of the advantages of delivering health interventions online is they can reach a large number of people relatively cheaply, and at a reduced cost compared to other approaches [2,26]. However, although hosting pages on social networking sites is free, the time spent creating, developing and maintaining them isn't. The time to upload posts can be substantial when multiple sites need to be updated, and posts monitored and responded

to. Sourcing and developing the content of posts also requires resources; even if sites are largely reliant on existing content, this must be sourced and reviewed for accuracy and appeal. We used an 'edutainment' (education and entertainment) approach to maximize appeal to our target audience, which required substantial investment to develop.

As our project involved a novel approach, we were unsure at the outset of the resources required. As the project evolved, we realized we had substantially under-estimated the time and effort required to develop the sites initially, and to maintain them for the duration of the project. Given the amount of information on social networking sites, and the speed at which information changes (on Facebook alone 30 billion pieces of content are shared each month [14]), we potentially needed to be posting content several times a day, rather than every few days. We found that the resources required to maintain sites detracted from attending to other key tasks, such as exploring alternative approaches to engage end users, maintaining our collaboration, and project evaluation. Upon reflection, it would have been ideal to have had the capacity to employ an individual with the time and interest to maintain the pages (eg, an avid social media user), rather than using a combination of creative professionals and researchers, whose primary project roles were not online maintenance.

Recommendation 4. Generate Interest (buzz) and Do It Early—Just Because You've Built It, Doesn't Mean They'll Come.

One of the greatest challenges for health promotion interventions using social networking sites is being noticed amongst the huge amount of content online. Unlike traditional advertising, being visually appealing is not sufficient to attract attention. It helps to have an established base of end users when the site is launched; feedback from our initial IT laboratory testing with end users indicated sites need to look active to attract interest from others. We attempted to do this by 'soft launching' our pages via word-of-mouth through personal and professional networks. However, this approach risked having an initial fan base different to the target demographic, which may limit the appeal of the intervention to the intended audience.

Promotion of the intervention is also critical; while we utilized 'traditional' methods of promotion such as print and broadcast media coverage and advertising, by far the most successful was using Facebook advertisements (although ours had an incentive attached) and uploading and tagging photos of end users at public events. Others have also noted the success of using online advertisements and photo tagging to attract end users [22,27]. Resources for promotion are most effectively spent in online strategies that allow end users to immediately 'click through' to sites, rather than a two-step process of viewing the advertisement and finding the site online. Having a defined offline community to reach (as we did for the MSM arm of the project) also assists with targeting promotion.

Recommendation 5. Keep Your Audience Engaged—Don't Fall Off the Newsfeed!

Users frequently connect with pages and groups, and download applications, never to take notice of them again. The amount of content available is overwhelming; Facebook alone has 900 million pages, groups, events and community pages [14]. This presents a difficulty for the delivery of health promotion online, especially when sustained engagement over time is required to deliver the intervention.

We were conscious at the outset that we did not want to deliver a Web 1.0 intervention using a Web 2.0 site. We aimed to truly interact and engage with our target group, not just broadcast information. The challenge was to maintain interest and engagement over a four-month intervention period with sufficient audience reach. We attempted to do this by using different delivery mechanisms such as posting regular updates (both text and videos), posing questions and encouraging comments on posts, and launching quizzes and polls, with varying success. However, it was clear from site usage data that interest and interaction on our pages declined considerably over time in both project arms (Figure 3 and Figure 4). In addition, the use of multiple delivery mechanisms may have 'fragmented' our key health messages; even if an individual had been exposed to one delivery mechanism, they may not have received the full message if they did not view other content on the site.

The loss of participants over time within an online intervention – the 'Law of Attrition' – is well known, and is often simply due to loss of interest of participants [10,28]. In retrospect, we may have been able to increase (or simply maintain) engagement by delivering the intervention over a shorter time frame, focusing on a single core message, ensuring all posts could act as 'stand alone' messages and creating more opportunities for end users to generate and manipulate content themselves. Further investigations are needed to establish the optimum methods to engage users of social networking sites in health promotion interventions, and to retain them over time.

Recommendation 6. Go Viral—If You Find the Formula, You're a Millionaire.

Ultimately, the greatest success one can have on a social networking site is "going viral" —where enough people are sufficiently interested in a post to share it with their friends, who then share it with their friends and so on, resulting in an exponential growth of connections. This spread of information has been termed 'Internet meme' [29]; the most common examples are when videos go viral and attract millions of views (eg, "Dancing Matt", "Obama Girl", "Diet Coke + Mentos" [30]). Even some videos containing health-related content have managed to achieve this; for example, "Kicesie's Sex Ed" YouTube channel has attracted over 240 million views [31]. In terms of health promotion, the aim would be to achieve viral spread primarily within the target population, as a widely dispersed intervention may be of little value if it does not reach the intended audience.

The challenge for those developing interventions on social networking sites is that no formula exists for achieving viral spread, and we certainly didn't achieve this with our project (our most popular video had 3118 views). The critical factors believed to be important for viral spread include the structure of the campaign (if it is structured to encourage viral activity, and if it complies to ethical standards), the product being marketed by the campaign (if it is suitable for viral spread) and

the message content (if the message is imaginative, contains fun and intrigue, is accessible and is engaging) [32]. Others stress the importance of having individuals with exceptionally large numbers of social connections to share the message [33]. Available empirical data supports that positive content, content inspiring emotion (particularly awe and surprise), content capturing imagination, and the connectivity of the person transmitting the information are important in ensuring viral spread [34-37]. Additionally, viral spread alone may not be enough: while it may increase viewing of one piece of content, this may not translate into sustained interest and engagement. Currently, our best suggestion is to keep trialling different strategies targeted to your audience; hopefully you'll be lucky and hit the jackpot!

Recommendation 7. Define Success and How You Will Measure It

There is little point developing health promotion interventions on social networking sites if it is not possible to measure if they are successful. This brings about two challenges: how to define success and how best to measure it.

As our project was a pilot we had both 'process' and 'impact' evaluation aims. These evaluation aims included assessing whether we could develop an intervention on social networking sites and attract and engage end users whilst delivering health promotion messages that would have a positive effect on sexual health knowledge and behaviour. As with any approach in its infancy, it is appropriate to focus on process as well as impact evaluation outcomes [38].

An appropriate methodology is of critical importance when evaluating interventions on social networking sites. Not only may we wish to evaluate traditional process and impact outcomes for health interventions (eg, reach, dose delivered and received, knowledge and behavioural changes) [38,39], the usability and appeal of the sites is also of key importance. Evaluations of interventions using social networking sites need to appropriately define and measure end user engagement, and develop ways of measuring if and how engagement assists with achieving intervention aims; for example, is a 'like' of a page a valid measure of engagement, or is only a comment indicative of true user engagement. Evaluation in this setting is complicated further by the fragmenting of health messages across delivery mechanisms; it can be complex to measure which messages and delivery mechanisms end users were exposed to, and whether this exposure translated into any degree of positive behaviour change.

For our project, we integrated evaluation methods derived from both the health (questionnaires, focus groups, diaries) and information technology (user laboratory testing, expert review) spheres. This is consistent with O'Grady's proposed 'dynamic framework' that suggests incorporating technology (eg, system robustness, reliability, usage statistics) and computer-mediated interaction (eg, usability, accessibility, interactivity) elements within system evaluations [40]. To establish the evidence base for how best to use social networking sites for health promotion interventions, it is critical to move beyond simply collecting end user statistics and integrate evaluation methods from multiple disciplines.

Conclusion

Although there is much discussion and interest about using social media for health promotion interventions [10,16,17], our experience suggests this is far easier said than done, particularly if the intervention aims to truly use Web 2.0 functions to engage end users. Developing an intervention on social networking sites requires consideration of additional aspects beyond more traditional methods of health promotion. Developers need to consider the online environment and the nature of human interaction online, including Web 2.0 functionality, the characteristics of the target audience and their preferred social networking site(s), and how end users interact and engage in these spaces. Additionally, obtaining ethical, legal and organizational approval, and developing effective evaluation strategies may be challenging. These aspects require additional expertise not typically found in health-focused organizations, and the investment of considerable time and resources.

Social networking sites are now an established part of the online environment; despite being less than ten years old, they are among the most frequently accessed sites globally [13]. While the particular site that is most popular may change over time [12], these sites share common functions that have fundamentally changed how individuals communicate and interact both on- and off-line. Although these sites are primarily used to communicate with social networks, the increasing amount of time individuals spend in these settings [41] suggests that health organizations need to develop effective strategies for reaching individuals in these spaces, whether delivering interventions or using these sites to promote interventions delivered elsewhere.

The FaceSpace Project was our first attempt to develop a health promotion intervention using social networking sites. At the time of project conception there was no information in the published health literature to guide our project development, and undoubtedly we made several mistakes throughout the process. However, our staged implementation approach ensured we could incorporate learnings from the first arm into the second (and now into the extension of the MSM arm), and we were able to develop an appropriate evaluation strategy.

As the popularity of social networking sites continues to increase, we hope that our experience is able to inform the development and evaluation of future health promotion interventions in these spaces. Developing health promotion interventions in this setting, and making mistakes and learning from them is certainly far better than doing nothing at all [30]. With the continuing change in communications media, health organizations must embrace these technologies or risk being left behind.

Acknowledgments

Judy Gold receives funding from the Australian Government through an Australian Postgraduate Award, and a Monash University Faculty of Medicine Excellence and Postgraduate Publication Award. Alisa Pedrana receives funding from the Australia Government through a National Health and Medical Research Council (NH&MRC) Public Health Postgraduate Scholarship, and the Sidney Myer Health Scholarship. Margaret Hellard receives funding from the NH&MRC as a senior research fellow. This project was funded by the Victorian Department of Health. The funders did not play any role in study design or conduct or manuscript preparation.

Authors' Contributions

All authors were involved in the design and implementation of The FaceSpace Project. JG led the conception, preparation and review of this manuscript. AP and MS were involved in manuscript conception and preparation. SH, SC, JA, OL and CB contributed to the recommendations presented in the manuscript, and reviewed the manuscript. MH was involved in manuscript conception and preparation.

Multimedia Appendix 1

Content Monitoring Protocol.

[PDF File (Adobe PDF File), 28KB - jmir_v14i1e30_app1.pdf]

References

1. World Development Indicators. 2010. Internet users (global) URL: http://data.worldbank.org/indicator/IT.NET.USER [accessed 2012-02-13] [WebCite Cache ID 65Q8E33tt]
2. Strecher V. Top 500 Global Sites. Journal of Computer-Mediated Communication 2011;3:33-45. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]
3. Fox S. The Social Life of Health Information. Washington, DC: Pew Internet & American Life Project 2009 URL: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2011-06-28] [WebCite Cache ID 5zmDS2IBq]
4. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. Cochrane Database Syst Rev 2010(9):CD006483. [doi: 10.1002/14651858.CD006483.pub2] [Medline: 20824850]
5. Civljak M, Sheikh A, Stead LF, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2010(9):CD007078. [doi: 10.1002/14651858.CD007078.pub3] [Medline: 20824856]
6. 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]
7. 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]
8. O'Reilly T. O'Reilly. 2005. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software URL: http://www.oreillynet.com/lpt/a/6228 [accessed 2012-02-13] [WebCite Cache ID 65Q8VWYAS]
9. O'Reilly T, Battelle J. Web 2.0 summit. 2009. Web Squared: Web 2.0 Five Years On URL: http://assets.en.oreilly.com/1/event/28/web2009_websquared-whitepaper.pdf [accessed 2011-06-28] [WebCite Cache ID 5zmDvkQvP]
10. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr 29;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
11. Safko L, Break DK. The Social Media Bible: Tactics, Tools, and Strategies for Business Success. Hoboken, New Jersey: Wiley; 2009.
12. Boyd D, Ellison N. Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication 2007;13(1):210-230 [FREE Full text] [doi: 10.1111/j.1083-6101.2007.00393.x]
13. Top 500 Global Sites. Alexa. 2011 URL: http://www.alexa.com/topsites/global [accessed 2011-06-28] [WebCite Cache ID 5zmEIPxhs]
14. Facebook. 2010. Facebook Statistics URL: http://www.facebook.com/press/info.php?statistics [accessed 2011-06-28] [WebCite Cache ID 5zmEKyeRi]
15. Zuckerberg M. The Facebook Blog. 2009. 200 Million Strong URL: http://blog.facebook.com/blog.php?post=72353897130 [accessed 2011-06-28] [WebCite Cache ID 5zmENcZsp]
16. Freeman B, Chapman S. Gone viral? Heard the buzz? A guide for public health practitioners and researchers on how Web 2.0 can subvert advertising restrictions and spread health information. J Epidemiol Community Health 2008 Sep;62(9):778-782. [doi: 10.1136/jech.2008.073759] [Medline: 18701726]
17. Rietmeijer CA, McFarlane M. Web 2.0 and beyond: risks for sexually transmitted infections and opportunities for prevention. Curr Opin Infect Dis 2009 Feb;22(1):67-71. [doi: 10.1097/QCO.0b013e328320a871] [Medline: 19532082]

18. Bardus M, Suggs L. Social media & social marketing: Are we there yet? 2010 Presented at: 20th IUHPE World Conference on Health Promotion; 11-15 July, 2010; Geneva, Switzerland.
19. Berglas N, Ralph L, Schwartz S, Brindis C. Finding Youth in Their Space: Using Social Networking Sites (SNS) to Connect Youth to Sexual Health Services. San Francisco: Sex::Tech Conference 2009; 2009 Presented at: Sex:Tech Conference; March 22-23, 2009; San Francisco, CA.
20. Folsom A, Jackson J. From the Internet to Action: Combining Social Marketing with Media to Empower Teens to Take Control and Get Tested. 2009 Presented at: Sex::Tech Conference; March 22-23, 2009; San Francisco, CA.
21. Hoskin T, Hounslow S. Social Networking Sites and Tobacco-Free Education: Do They Mix? 2008 Presented at: 7th Annual Australian and New Zealand Adolescent Health Conference; November 5-7, 2008; Melbourne, Australia.
22. Schmidt M, Currie M, Bertarm S, Bavinton T, Bowden FJ. Promoting Chlamydia Testing to Young Women, Their Partners and Their GPs Using Modern Social Media. 2009 Presented at: Australasian Sexual Health Conference; September 7-9, 2009; Brisbane, Australia.
23. Moreno MA, Briner LR, Williams A, Brockman L, Walker L, Christakis DA. A content analysis of displayed alcohol references on a social networking web site. J Adolesc Health 2010 Aug;47(2):168-175. [doi: 10.1016/j.jadohealth.2010.01.001] [Medline: 20638009]
24. Moreno MA, Vanderstoep A, Parks MR, Zimmerman FJ, Kurth A, Christakis DA. Reducing at-risk adolescents' display of risk behavior on a social networking web site: a randomized controlled pilot intervention trial. Arch Pediatr Adolesc Med 2009 Jan;163(1):35-41 [FREE Full text] [doi: 10.1001/archpediatrics.2008.502] [Medline: 19124701]
25. Ahern DK. Challenges and opportunities of eHealth research. Am J Prev Med 2007 May;32(5 Suppl):S75-S82. [doi: 10.1016/j.amepre.2007.01.016] [Medline: 17466822]
26. 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]
27. Hoppe M. Successful Online Survey Recruitment Strategies. San Francisco, CA; 2010 Presented at: Sex::Tech Conference; February 26-27, 2010; San Francisco, CA.
28. 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]
29. Marshall G. The Internet and Memetics. 1998 Presented at: 15th International Congress on Cybernetics Symposium on Memetics; August 24-28, 1998; Namur, Belgium URL: http://pespmc1.vub.ac.be/conf/memepap/marshall.html
30. Qualman E. Socialnomics: How social media transforms the way we live and do business. Hoboken, New Jersey: Wiley; 2009.
31. Kicesie's Sex Ed. URL: http://www.youtube.com/user/kicesie [accessed 2011-06-28] [WebCite Cache ID 5zmEswDcK]
32. Woerndl M, Papagiannidis S, Bourlakis M, Li F. Internet-induced marketing techniques: Critical factors in viral marketing campaigns. Int J of Business Science and Applied Management 2008;3(1):33-45.
33. Kaplan AM, Haenlein M. Two hearts in three-quarter time: How to waltz the social media/viral marketing dance. Business Horizons 2011;54(3):253-263.
34. Berger JA, Milkman KL. Social Transmission, Emotion, and the Virality of Online Content. SSRN eLibrary 2009 [FREE Full text]
35. Dobele A, Lindgreen A, Beverland M, Vanhamme J, van Wijk R. Why pass on viral messages? Because they connect emotionally. Business Horizons 2007;50(4):291-304.
36. Stephen AT, Dover Y, Goldenberg J. A Comparison of the Effects of Transmitter Activity and Connectivity on the Diffusion of Information Over Online Social Networks. INSEAD Working Paper No. 2010/35/MKT. SSRN eLibrary 2010 [FREE Full text]
37. De Bruyn A, Lilien GL. A Multi-Stage Model of Word of Mouth through Viral Marketing. Int J of Marketing 2008;25(3):143-225.
38. Nutbeam D, Bauman A. Evaluation in a nutshell: a practical guide to the evaluation of health promotion programs. North Ryde, Australia: McGraw-Hill; 2006.
39. Steckler A, Linnan L. Process evaluation for public health interventions and research. San Francisco, CA: Jossey-Bass; 2002.
40. O'Grady L, Witteman H, Bender JL, Urowitz S, Wiljer D, Jadad AR. Measuring the impact of a moving target: towards a dynamic framework for evaluating collaborative adaptive interactive technologies. J Med Internet Res 2009;11(2):e20 [FREE Full text] [doi: 10.2196/jmir.1058] [Medline: 19632973]
41. Nielsen. 2009. Global Faces and Networked Places: A Nielsen report on Social Networking's New Global Footprint URL: http://blog.nielsen.com/nielsenwire/global/social-networking-new-global-footprint/ [accessed 2011-06-28] [WebCite Cache ID 5zmFacHcF]

Abbreviations

MSM: Men who have sex with men

Edited by G Eysenbach; submitted 29.06.11; peer-reviewed by C Fry, S Sarbadhikari, T Biswas; comments to author 02.09.11; revised version received 24.11.11; accepted 14.01.12; published 28.02.12.

Please cite as:
Gold J, Pedrana AE, Stoove MA, Chang S, Howard S, Asselin J, Ilic O, Batrouney C, Hellard ME
Developing Health Promotion Interventions on Social Networking Sites: Recommendations from The FaceSpace Project
J Med Internet Res 2012;14(1):e30
URL: http://www.jmir.org/2012/1/e30/
doi:10.2196/jmir.1875
PMID:22374589

Original Paper

What Are Young Adults Saying About Mental Health? An Analysis of Internet Blogs

Madalyn A Marcus[1], MA; Henny A Westra[1], PhD; John D Eastwood[1], PhD; Kirsten L Barnes[1], BA(Hons); Mobilizing Minds Research Group[2]

[1]Department of Psychology, York University, Toronto, ON, Canada

[2]see acknowledgements

Corresponding Author:
Madalyn A Marcus, MA
Department of Psychology
York University
4700 Keele Street
Toronto, ON, M3J 1P3
Canada
Phone: 1 416 736 2100 ext 33706
Fax: 1 416 736 5814
Email: madalyn@yorku.ca

Abstract

Background: Despite the high prevalence of mental health concerns, few young adults access treatment. While much research has focused on understanding the barriers to service access, few studies have explored unbiased accounts of the experiences of young adults with mental health concerns. It is through hearing these experiences and gaining an in-depth understanding of what is being said by young adults that improvements can be made to interventions focused on increasing access to care.

Objective: To move beyond past research by using an innovative qualitative research method of analyzing the blogs of young adults (18–25 years of age) with mental health concerns to understand their experiences.

Methods: We used an enhanced Internet search vehicle, DEVONagent, to extract Internet blogs using primary keywords related to mental health. Blogs (N = 8) were selected based on age of authors (18–25 years), gender, relevance to mental health, and recency of the entries. Blogs excerpts were analyzed using a combination of grounded theory and consensual qualitative research methods.

Results: Two core categories emerged from the qualitative analysis of the bloggers accounts: I am powerless (intrapersonal) and I am utterly alone (interpersonal). Overall, the young adult bloggers expressed significant feelings of powerlessness as a result of their mental health concerns and simultaneously felt a profound sense of loneliness, alienation, and lack of connection with others.

Conclusions: The present study suggests that one reason young adults do not seek care might be that they view the mental health system negatively and feel disconnected from these services. To decrease young adults' sense of powerlessness and isolation, efforts should focus on creating and developing resources and services that allow young adults to feel connected and empowered. Through an understanding of the experiences of young adults with mental health problems, and their experiences of and attitudes toward receiving care, we provide some recommendations for improving receptivity and knowledge of mental health care services.

(J Med Internet Res 2012;14(1):e17) doi:10.2196/jmir.1868

KEYWORDS

Young adult; mental health; mental health services; life experiences; blogging; qualitative research

Introduction

Mental health problems are highly prevalent among young adults, with up to 25% experiencing a mental health problem in a given year [1-3]. Despite this high prevalence, young adults are particularly unlikely to seek help and, as a result, many do not receive adequate care [4-6]. Research has shown that stigma and negative beliefs about mental health care play a fundamental

role in the decision to access and remain engaged in care [5,7-9]. Thus, to understand why young adults are particularly unlikely to access treatment, it is important to determine the specific beliefs and experiences of young adults with mental illness. Given the tendency of young adults to avoid seeking help, innovative ways of hearing from these hard-to-reach young adults must be explored. The Internet provides such a possibility.

Going Online

The Internet has become a key space for health information sharing [10], with 78% of American young adults looking for health information online [11]. In fact, people are more likely to use the Internet to find health information than to go to a physician [12]. Moreover, the Internet is not solely a location to gather health information, as people frequently use online spaces to tell their stories and connect with others. This is especially the case with online journals or blogs, which are updateable public records of private thoughts [13]. Previously termed weblogs, blogs gained popularity in 1999 by providing individuals with the ability to write about their thoughts and feelings in a free, dedicated online space [14]. BlogPulse [15], a trend-discovery system for blogs, monitors over 155 million blogs with over 1 million updated in the last 24 hours. In 2010, 12% of Internet users wrote blogs, while 51% of Internet users read blogs [16]. Rates vary on the frequency of blog usage by young adults, with 18% to 39% of young adults reporting having written an online journal or blog [17,18]. Further the number of blogs and blog writers is growing exponentially [19,20]. Thus, the widespread use of the Internet has created a unique space for hearing from young people who may not have been consulted in previous health service utilization research.

Despite the large number of young adults using the Internet to share their experiences, few studies have examined these accounts to gain an increased understanding of the experiences of young adults with mental health problems. Rather, most Internet-based studies actively recruit participants using methods such as online surveys [21] or qualitative interviews [22-24]. Baker and Fortune [22] found through email-based semistructured interviews with young adults that an online self-harm or suicide community provides emotional support, valuable information, and friendship for individuals. While these studies provide vital information, they do so within the confines of a research study where participants are recruited, and thus their responses are influenced by participants' awareness of the research context. This contextual knowledge introduces demand characteristics, which have been found to significantly influence participant responses [25].

Present Study

The present study used an innovative approach to understand the overall experience and impact of living with a mental health problem among young adults. Namely, we conducted a qualitative grounded theory analysis of the blogs of young adults (18–25 years of age) who were specifically blogging about their experiences with mental health problems. As such, the present study allowed young adults' experiences to be understood free from a research context, thereby providing a more unbiased account of the experience of living with a mental health problem.

Methods

Selection of Blogs

We used an enhanced search engine, DEVONagent (Mac OSx; DEVONtechnologies, LLC, Coeur d'Alene, ID, USA), to search online blogs with English text that included key words pertaining to mental health. Initial searches using keywords such as depression yielded a high volume of possible blogs. We subsequently iteratively refined searches by using keywords identified in pertinent blogs. Multiple searches were conducted using various combinations of the following key words: blog (diary, personal, personal experience), mental health (depressed, depression, anxious, anxiety, bipolar), and therapy (counseling, support, psychologist, psychiatrist, medication, Prozac, Celexa, etc). Later searches added the keywords youth and young adulthood. Subsequently, we manually searched the blogs from the initial searches. That is, we then reviewed each bloggers' blog roll, which is a list of other blogs that the current blogger recommends. This allowed us to connect to other blogs that were also affiliated with mental health issues in order to identify other possible bloggers and major websites (eg, Anxiety Tribe, Depression Tribe, and PsychCentral) meeting criteria relevant to the study. Thus, we downloaded an exhaustive list of existing personal online and publicly available blogs on the theme of mental health issues and archived them onto a secure computer.

This generated a large database (approximately 3500 webpages), which was furthered refined in the following steps. First, we included only blogs from those 18 to 25 years of age. Blogs were excluded when no age was listed in the user profile or age could not be determined from the text of the blog. Second, we checked the blogs to ensure that the content of the blog pertained primarily to young adult discussion of mental health problems. Third, to ensure that selected blogs were recent and frequently updated, only blogs posted between August 2008 and February 2009 were included. Fourth, only blogs that were updated on a weekly or biweekly basis were retained. Fifth, we examined the frequency of viewing of the blog by others and removed blogs that had been viewed less than 200 times. At this stage in the data selection procedure, we had reduced our sample to 18 blogs, the majority of which were authored by women with mood disorders. To ensure more proportional representation of men and a greater variety of mental health concerns, we eliminated another 10 blogs from the sample that were authored by women and focused on mood disorders. Sixth, we retained excerpts only from the blogs pertaining to mental health for analysis and excluded irrelevant content (eg, events or discussion unrelated to mental health). In general, extraneous material not related to mental health was relatively rare, comprising less than 5% of the blog postings.

Participants

In summary, our selection procedure resulted in a total of 8 blogs with the following characteristics: clearly authored by 18- to 25-year-olds; frequently updated and viewed; authored by 2 men and 6 women who had a variety of mental health problems; and written between August 2008 and February 2009.

Table 1. Sample characteristics of bloggers

Blogger	Gender	Age (years)	Location	Main mental health concern
1	Female	23	England	Bipolar disorder
2	Male	23	England	Social phobia
3	Male	22	Ireland	Bipolar disorder
4	Female	23	United States	Posttraumatic stress disorder/dissociative identity disorder
5	Female	21	England	Depression
6	Female	18	England	Depression/eating disorder
7	Female	22	Columbia	Depression
8	Female	18	England	Social anxiety

Table 2. Characteristics of blogs

Blogger	Total number of posts during time period	Number of text pages analyzed	Number of hits (at time of data collection)	Blog URL
1	130	30	500,000 hits since February 2007	http://thesecretlifeofamanicdepressive.wordpress.com/
2	51	29	1000 hits since August 2007	http://socialphobic.co.uk/
3	26	20	400 hits since May 2008	http://guyinterrupted.blogspot.com
4	73	12	231 hits since June 2008	http://crackersandjuice.blogspot.com/
5	59	29	7000 hits since May 2008	Not available
6	108	69	3000 hits since September 2008	http://blueskiesandgreengrass.wordpress.com/
7	50	42	36,000 hits since October 2006	http://crazyasuka.wordpress.com/
8	27	13	2,600 hits since September 2008	http://meryine.blogspot.com/

Table 1 presents characteristics of the final sample of 8 bloggers in this study. Table 2 presents information on the number of posts and hits for each blogger, the number of pages of text analyzed, and the blog URL. Across the 8 bloggers a total of 524 blog posts, comprising 244 pages of text, were reviewed and analyzed.

Ethical Issues

This study was reviewed and approved by York University's Institutional Review Board for research with human subjects. Only publicly accessible blogs were used in the present study. We excluded blogs requiring a username and password or registration form or fee for which the individual could clearly expect anonymity. The use of public rather than private information is consistent with recommendations for ethical conduct of research described by investigators conducting similar research and follows ethical conduct for Internet research in particular [26-28]. Further, the bloggers were individually contacted to obtain their consent for inclusion of quotes from their blogs and the URL for their blog in the presentation of the findings from the present study. Finally, all personally identifiable information was removed or changed.

Data Analysis

We analyzed blog excerpts using a combination of grounded theory [29] and consensual qualitative research methods [30]. Qualitative methods offer a systematic, inductive way of investigating experience. Rather than being constrained by previously determined concepts, they allow researchers to examine experiences and phenomena as they naturally arise from the data. More specifically, grounded theory is a method of analysis that emphasizes the generation of theory that is grounded in the inductive investigation of participant experiences, in this case, blogger accounts of their experience of mental health problems [29,31,32].

In grounded theory, researchers first immerse themselves in the text by reading and rereading the text as a whole. After this the text is divided into units of meaning reflecting a complete thought, which stays as close to the language of participants as possible [32]. This is referred to as open coding. Similar-meaning units are then continuously grouped together as coding proceeds into larger-order categories, both within and across participants. Categories are continuously modified throughout the analysis by adding and subtracting meaning units. This is referred to as the constant comparative method [29], where data and the conceptualized categories are continuously compared. This process eventually results in a hierarchy of categories, with lower-order categories being subsumed by higher-order ones. Often, an overarching core category (or categories) emerges that organizes the relationship among all other categories and the data supporting them.

Grounded theory typically relies on a single researcher. In the present study, however, consensual qualitative methods supplemented the grounded theory analysis to integrate the perspectives of multiple researchers. In consensual qualitative research a team of researchers is used to make decisions on the data by consensus [30]. The team consisted of 3 researchers, with 2 coding each transcript, and the third serving as an auditor. In the present study, meaning units and categories were coded independently by 2 researchers and then a final designation of meaning units and categories was determined through discussion to achieve consensus. The 2 coders and the auditor came together to develop consensus on the meaning units and categorization of the data. A fourth person served as an additional auditor at the end of this process for consensus on the overall model.

To achieve diversity in perspective on the data, the 2 coders and the 2 auditors differed in level of experience in treating mental health problems, with 1 senior undergraduate psychology student, 1 advanced graduate student in clinical psychology, and 2 clinical psychologists with 10 and 16 years of experience in the field, respectively. It is important for qualitative researchers to be aware of subjective biases and, as much as possible, to put aside or bracket these biases, expectations, and hypotheses. As such, in an effort to mange their assumptions, the researchers wrote field notes and memos, bracketing and becoming aware of any biases or personal reactions to the data [33] in order to remain as objective as possible. Through writing memos and self-reflection, the researchers, at all points throughout the study design, data collection, and data analyses, attempted to exhibit reflexivity by constantly examining how the research process, including their potential biases, may affect the resulting research outcomes [34].

Saturation

Saturation is the point at which the addition of new data does not add new information to the developing heuristic model. We considered the issue of saturation to be relevant both within and across bloggers. Within bloggers, meaning units were identified beginning with the most recent post and moving backward in time. During this process similar meaning units were placed together in categories. The coders independently identified the point at which no new categories were derived from reviewing additional blog posts. Coders then arrived at consensus through discussion, to identify the point of saturation across bloggers. In the present study, saturation across bloggers was achieved with 6 bloggers, since the addition of the last 2 bloggers did not yield any additional categories to the developing heuristic model.

Results

Two core categories emerged from analysis of the bloggers accounts: *I am powerless* (intrapersonal) and *I am utterly alone* (interpersonal). Table 3 presents the number of bloggers who wrote content pertaining to a given theme and the number of instances of a given theme across all bloggers. The emergent model of bloggers' experience of a mental health concern is depicted in Figure 1. The core categories and the themes (or categories and subcategories) they comprise are described further below.

Figure 1. Emergent model of bloggers' experience of a mental health concern.

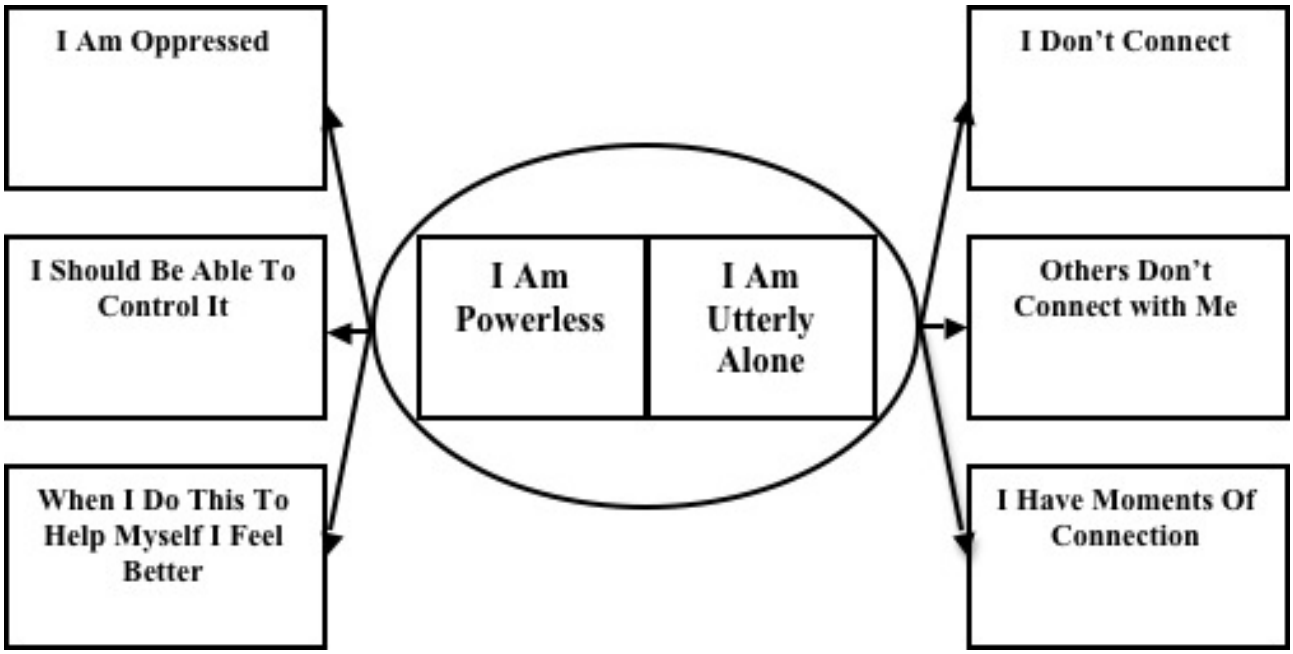


Table 3. Model of bloggers' experience of a mental health concern

Category	Number of bloggers	Number of meaning units across all bloggers
Core c ategory I: I am powerless		
I am oppressed		
It dominates my life/stains everything	8	31
It's relentless	7	32
It paralyzes me	8	19
It's exhausting	5	10
It's confusing/I don't know what to do	8	24
It makes me afraid	7	27
It's hopeless	8	19
I think about suicide	5	11
I should be able to control it	7	33
When I do things to help myself I feel better		
Self-care	7	29
Taking medications	4	14
Blogging	4	10
Core c ategory II: I am utterly alone		
I don't connect		
I hide my true feelings:	7	16
Because I'm ashamed	5	13
Because I'm concerned about others' reactions	6	16
Because I'm scared/unsure	3	16
I'm all alone	5	13
I lack the skill to connect	4	18
I am a burden	4	13
It's my fault that I'm alone/I deserve it	7	28
Others don't connect with me		
Others insist I have control when I don't	8	30
Mental health professionals are unresponsive		
I have strong mixed feelings about medication	7	46
I'm not getting enough care/I'm abandoned	6	42
My care is inconsistent/disorganized	2	13
Mental health professionals are unapproachable		
They are unsupportive, hurtful, or intimidating	6	18
They have their own agenda	5	10
Previous lack of support contributed to my problem	3	18
I have moments of connection:		
Because of me	8	30
Because of others	7	25
Because of blogging		
Blogging connects me to others who understand	6	14
Blogging allows me to help others	3	8

Core Category I: I am Powerless

Overall, young adult bloggers expressed significant feelings of powerlessness as a result of their mental health concerns. They described their mental health problems as an all-encompassing, highly destructive force, leaving virtually no aspect of their functioning and daily lives untouched. Young adults described feeling victimized and overpowered by their mental health problems, while simultaneously blaming themselves for not being able to control these experiences. They also articulated trying to engage in self-care activities, which provided some relief.

I am Oppressed

This category was defined by experiences of mental health problems as dominating, relentless, paralyzing, exhausting, confusing, frightening, and hopeless. At least 5 of the 8 bloggers contributed to each subcategory in the larger category of feeling oppressed, reflecting the commonality of these experiences across individuals with varied mental health problems.

It Dominates My Life/Stains Everything

All bloggers (8/8) contributed properties to this subcategory, describing mental health problems as permeating all aspects of their lives. This was often accompanied by strong feelings of frustration and resentment. For example, bloggers noted that

> *The biggest causalities have been my relationships with lovers, friends, family, my working life, my education, and my memory.* [blogger 1]
>
> *[My] moods go from low to high, and vice versa, continually...it's ruining my life.* [blogger 6]
>
> *I'm not all about depression but it takes so much of me. It doesn't define me but it stains everything I say and do.* [blogger 7]

It's Relentless

The majority of the bloggers (7/8) described feeling chronically incapacitated by the relentlessness of their mental health symptoms, including the significant fear of symptom relapse. They reported experiencing disappointment and frustration at the recurring nature of their symptoms. Bloggers stated that

> *It's 10 years of unchecked mental illness that has, over and over again, ruined whatever patchwork mockery of a life I had.* [blogger 1]
>
> *I don't want to be like this. I thought I was recovering. I thought that this was over. But it really isn't.* [blogger 6]

It Paralyzes Me

All bloggers (8/8) also described feeling restrained and controlled by their symptoms. For example, blogger 1 noted that

> *It's like being locked in an iron maiden...[Depression] pushes one into the rut of not having the energy or inclination for very much at all, but feeling resentful and frustrated because of it.*

And blogger 7 described that

> *Having depression is like having invisible cuffs. You don't move and you shift between feeling like the idiot who can't move although they have no restrains at all, and realizing the cuffs are indeed there but not being able to get rid of them.*

It's Exhausting

Relatedly, many of the bloggers (5/8) described the experience of having a mental health problem as exhausting. They spoke of the emotional and physical fatigue that was a constant aspect of their mental health problems. Bloggers also discussed their sleep difficulties:

> *[My mood] is stable, apart from irritability that is mostly due to sleep deprivation.* [blogger 7]
>
> *There is nothing worse than having a restless sleep when you need about 10 hours of it to keep you going and keep you sane.* [blogger 8]

It's Confusing/I Don't Know What to Do

All of the bloggers (8/8) described feeling perplexed by the symptoms themselves and about how to address them. For example:

> *I am stuck. I don't know if I can get better...and I don't seem to have the words to say what it is that I am going through. I cannot define my pain or explain it, I just have symptoms of it.* [blogger 4]
>
> *I don't really know what I want or need, but something has to change.* [blogger 4]

It Makes Me Afraid

The majority of bloggers (7/8) also discussed how they feared the power their mental health problem had over them. As a result, they frequently expressed a need to be "on guard" and always aware of their symptoms. Blogger 1 noted, "It all happens so astonishingly quickly I have to be so aware," and blogger 7 said:

> *Having atypical depression, my moods are relatively reactive. It becomes increasingly tiring to see yourself feeling better in response to something while being aware that as soon as that thing is gone, you'll fall back. That leads to anxiety when something good is happening.*

It's Hopeless

All bloggers (8/8) expressed often profound pessimism and hopelessness about the future as a result of having a mental health problem. For example:

> *Every time I think about the future, I just can't help but feel dread. The only comfort I have is knowing that at least I have the necessary physical ability to kill myself.* [blogger 2]
>
> *There's really no hope for someone like me, I just cannot cope with the struggles and hard things life throws at you.* [blogger 2]
>
> *I feel like I'm being watched on stage. I'm wearing too much makeup, this isn't who I am, I hate myself. I want to give up, I can't stand this.* [blogger 8]

I Think About Suicide

Just over half of the bloggers (5/8) reported extreme hopelessness in the form of suicidal ideation and appeared to derive comfort from entertaining the possibility of suicide. Examples:

> *I still find the thought of suicide comforting.* [blogger 1]

> *I am pretty much resigned to ending my life now...I am too cowardly to slit my wrists...I'll try harder to find a place to jump from...There is not one consistent mode of thinking or conscious thoughts about jumping off a bridge, it's more of an overall feeling that comes in waves.* [blogger 2]

Bloggers also wrote about feeling their suicide would not affect anyone. For example, blogger 6 wrote

> *When I commit suicide people won't remember me. The way I want it to be, because at the end of the day I'm not special to anyone, and the more people who can forget me the better.*

I Should Be Able to Control It

Even though they felt highly victimized and experienced mental health problems as an oppressive force, the majority of the bloggers (7/8) believed they were to blame for their problems and should be able to control their symptoms. Bloggers blamed themselves for their current symptoms, as exemplified by blogger 7: "It is easier to blame the problem on something you can potentially control—yourself—than to blame it on external influences" and later wrote "You also wonder what the big deal is...You probably brought this on yourself" (blogger 7). Moreover, bloggers frequently expressed feeling guilty for not being able to control their symptoms. For example, blogger 2 wrote that "I'm honestly sick of my life, the constant guilt that I should be making more of the opportunities that I have been extremely lucky to have."

When I Do Things to Help Myself I Feel Better

Despite feelings of profound powerlessness, the majority of the bloggers (7/8) described engaging in activities resulting in symptom improvement including self-care activities, taking medications, and blogging itself. When the young adults took active steps to decrease their symptoms, they experienced improvements physically, emotionally, and mentally.

Self-care

The majority of the bloggers (7/8) reported the use of helpful self-care activities. For instance, blogger 3 reported that

> *My mood is better today, less grim, more focused...I meditate to calm and center my mind, clear my head. Classical music (Chopin) plays in the background.*

In addition to mental and physical activities, some bloggers wrote about changes in their thinking that were beginning to result in an improved mood. For example, blogger 8 said

> *I feel like a failure in a way, but the amount I've been through and personally achieved in my time off has been completely life changing. I'm a stronger, wiser, and more experienced person.*

Taking Medications

Half of the bloggers (4/8) discussed the helpful aspects of taking medication for their mental health problems. For example, "The Xanax is helping me a lot...I'm looking forward to the future rather than being paralyzed by fear at the thought of it" (blogger 3).

Blogging

Interestingly, half of the bloggers (4/8) stated that they used blogging as a means of coping, and found that disclosing and discussing their experiences with mental health problems online served many functions, including self-reflection, self-help, and acquiring much needed support. For example:

> *The one thing that keeps me in treatment, and writing, is pure interest...It's fascinating.* [blogger 1]

> *I write these blogs as a way to keep up with my own mood, to monitor my progress (or lack thereof)...To be honest, without the Internet this would be rather difficult, if not impossible.* [blogger 7]

> *I'd prefer to spend hours writing...My worries float away as every page is turned...it's like hiding in my word world.* [blogger 8]

Core Category II: I Am Utterly Alone

The young adult bloggers described a profound sense of loneliness, alienation, and lack of connection with others. Young adults reported hiding their feelings from others because they felt ashamed, were concerned about other's reactions, felt that they were a burden, and were scared or unsure about receiving mental health care, which contributed to their feelings of loneliness. They also spoke of feeling alone and as if they did not fit into the world around them. The bloggers wrote about the strain their mental health problem put onto others, further contributing to feelings of isolation. Many of the bloggers in fact blamed themselves for being alone, as they felt they lacked the skills to connect with others and that they deserved this isolation. Young adults also indicated that others, including mental health professionals, did not connect with them, as they were unresponsive and unapproachable. Despite this prevalent experience of loneliness, they also wrote about moments of connection when they took active steps to engage with others. In particular, blogging they viewed blogging as a positive space for discussing their mental health concerns, which served to empower the young adults.

I Don't Connect

I Hide My True Feelings

The majority of the bloggers (7/8) spoke about hiding their experiences of living with a mental health problem. For example, blogger 1 wrote:

> *In the light of someone speaking to me, I will smile and respond, but as soon as they turn their heads my face falls immediately.*

Similarly, blogger 3 wrote:

> *I put on a cheerful face, an easy smile. But it's all empty, it's all a façade. I'm so exhausted all I do is take my pills, eat, and sleep. I'm barely functioning.*

Young adults reported hiding their true feelings because of shame, concern about other's reactions, and uncertainty about these feelings (which are described below).

Because I'm Ashamed

Bloggers (5/8) reported hiding their feelings because they were ashamed about experiencing a mental health problem and assumed that others would share their negative self-view. This experience of shame and inferiority was often accompanied by fear of asking for assistance. For example, blogger 6 noted:

> *The day I have to tell the truth. The day where everything is going to come out, like a big ball of wool unraveling. How I don't eat, sleep and how I tried to throw myself off a bridge. Whilst making all this seem perfectly reasonable, and that I'm not really mental, honest. I know I need more help, I'm just scared of the consequences.*

These bloggers perceived themselves as weak and requiring more help to get though daily life than the average person, which contributed to their experience of shame. For example:

> *I had no choice but to ask for help. It was shaming for me. I hated myself for it.* [blogger 1]
>
> *I feel ridiculous requesting a softer treatment [help], and the first thing that suffers is my pride. But if I don't do something I'll drop out for sure.* [blogger 7]

Because I'm Concerned About Others' Reactions

Most bloggers (6/8) anticipated that others would not react supportively if they were honest about their experiences. For example, blogger 7 wrote

> *I almost never talk about the "real thing" outside this blog. People never react well even if they are well meant.*

Many of the young adults minimized the severity of their symptoms when speaking with others in an effort to protect people around them. Blogger 2 spoke about being concerned about her impact on her therapist:

> *It makes her sad to hear what I have to say...I don't really like it when she looks at me all concerned.*

Young adult bloggers spoke frequently about forcing themselves to appear "normal" for fear that others' negative reactions would make things worse for them. For example:

> *I do basic things trying to raise no suspicions [get myself food or put on clothes]...I don't want them to know how wrong things really are, because then they'll confront me, and I have no answers.* [blogger 7]

They also spoke of fear that others would reject them. For example:

> *[Wanting my therapist] to protect and comfort me is NOT ok, when you start wishing for them, people feel overwhelmed and suffocated, they perceive you as needy.* [blogger 4]

Because I'm Scared/Unsure

Some of the bloggers (3/8) wrote about feeling scared and ambivalent about receiving mental health care, which was another reason for hiding their feelings and not seeking help. For example:

> *I've emailed Samaritans...but they don't really help much and I'm far too scared to phone them.* [blogger 2]
>
> *I want to find out about more services in [the city I live in] and talk to the university counseling service but I don't have the confidence to.* [blogger 6]

I'm All Alone

Many of the bloggers (5/8) spoke about experiencing significant social challenges and feeling like outcasts from society. Examples of this profound sense of being alone and disconnected:

> *I don't fit in there either. It seems like there's no place for me in the world.* [blogger 2]
>
> *I feel disconnected from everything and everyone and I hate to pretend there's a connection anyway (it requires too much energy and it's futile).* [blogger 7]

This social isolation contributed to the negative experiences of living with a mental health problem.

I Lack the Skill to Connect

Repeatedly describing it as a "hard process," half of the bloggers (4/8) reported that they lacked the skill to connect with others. As blogger 2 reported, "It's hard being alone all the time. What little social skills I had have atrophied." The lack of connection with others contributed to their beliefs that they were unworthy of friendship and were incapable of forming social connections. Textbox 1 shows an excerpt from a poem blogger 8 wrote about this experience:

Textbox 1. Excerpt From a Poem Posted by Blogger 8

I know I'm a terrible friend,
It's a place I lack experience.
Always afraid to speak unless spoken to,
Wanting to run and hide.
Friends come along, get bored and move on.
For I've been hurt so many times,
Again I'm used.

Again I'm empty

Again I'm alone.

I Am a Burden

Half of the bloggers (4/8) wrote about how the strain their mental health problem put on family and friends further contributed to feelings of isolation. For example:

> *It's exhausting being depressed for the both of us [boyfriend and me].* [blogger 1]
>
> *My parents are already ashamed of me, the stupid overgrown child who can't handle life.* [blogger 2]

Consequently, young adults reported feeling guilty because of the negative impact their symptoms had on those around them.

It's My Fault That I'm Alone/I Deserve It

The majority of the bloggers (7/8) blamed themselves for their experience of loneliness. That is, they felt they somehow deserved their mental health symptoms and the resulting social isolation. For example, "I feel embarrassed and moronic, why should I inflict my worthless, helpless self on anyone?" (blogger 8), and "Everybody else manages fine, why not me?" (blogger 8).

Others Don't Connect With Me

Young adults described experiencing alienation through others' invalidating insistence that people with mental health problems should and can exert control over their experiences. Moreover, young adults also described finding mental health professionals unresponsive and unapproachable. The bloggers also attributed their current problems, at least in part, to previous experiences of lack of support and caring from others.

Others Insist I Have Control When I Don't

Consistently across all bloggers (8/8) was the expectation from others that their mental health problem should be easily manageable and, consequently, they experienced others as dismissive and offering simplistic advice. This is distinct from the intrapersonal feeling, discussed above, that one lacks control, because here bloggers are emphasizing the expectations of others. Textbox 2 shows a sample excerpt describing the frustration the bloggers feel as a result of this expectation.

Furthermore, the young adults wrote about being expected to just "get over it" and continue with regular activities of daily living. For example:

> *I'm constantly expected to perform tasks I struggle with. My family somehow forget that I've rarely ever been able to make phone calls, speak to people.* [blogger 8]
>
> *The attitude of my family is very telling. It is acceptable for me to feel awful because of something physical, as it is my body and I can't control it, right? But, mentally? *gasp* No way! I cannot feel awful, I have to be strong. Doing otherwise implies weakness and a failure as a person however physical this really feels.* [blogger 7]

Textbox 2. Excerpt From Blogger 1 Describing the Frustration With Others' Expectation That Mental Health Problems Are Easily Manageable

Some of you reading this will of course take the angle that I can't feel that bad because I'm sitting upright and writing this. Ah well. I feel guilty that I'm not cycling through the streets with hot cross buns in my basket like Mary Poppins. If someone were to whisper in my other ear that exercise cures depression I would probably punch them with the hand that I'm not using to chain smoke. I like to exercise, it makes me feel good but right now I simply don't have the energy. I would just veer straight into a wall. I do all that jazz you're supposed to do to "help yourself." I'd like someone to explain how positive thought helps when your mind specifically boycotts positive thoughts from entering the building.

Mental Health Professionals Are Unresponsive

I Have Strong Mixed Feelings About Medication

The majority of the bloggers (7/8) expressed substantive ambivalence toward medication as a result of prior negative experiences, such as side effects. For example

> *I hate, hate, hate taking medication. I honestly cannot decide if it is for the best or not.* [blogger 1]
>
> *Managing my illness my way didn't work...[Taking lithium] is the sentence I've been handed down. Let me think of it as freedom, and not my doom.* [blogger 3]

I'm Not Getting Enough Care/I'm Abandoned

The majority of young adult bloggers (6/8) reported feeling abandoned by the mental health care system, which resulted in a further sense of isolation and decline in functioning. Examples of this experience:

> *They can't help me except by cramming pills down my throat but that doesn't fix anything.* [blogger 2]
>
> *Tell him [the general practitioner] about everything...and I got a referral back to the primary care mental health team. I wasn't exactly pleased, because I knew that I needed a lot more than a waiting list and no therapy.* [blogger 6]

My Care is Inconsistent/Disorganized

Some of the bloggers (2/8) described their mental health care as incoherent, confusing, and unpredictable. Blogger 6 stated that "I'm just wondering if my referral will ever go through."

Mental Health Professionals Are Unapproachable

They Are Unsupportive/Hurtful/Intimidating

Many of the young adult bloggers (6/8) reported experiencing their psychiatrist or therapist, in addition to being neglectful, as intimidating, hurtful, and unsupportive. Blogger 7 expressed discontent toward her physician:

> *I'm tired of the psychiatrist and her dull look who just sits and writes things on her chart attempting one thing after the other like it was nothing for me to take one failure after another "hey! it didn't work! Let's try something else!*

Compounding this experience was the bloggers' described resistance to the therapist due to their lack of support. For example:

> *I get scared he's [the psychiatrist] going to get angry at what I say, especially after he smashed his fist on the table after I deigned to ask him about antianxiety medication.* [blogger 2]
>
> *The therapist really does not like my eating disorder. It seems to be the one thing she doesn't have a lot of compassion for and wants me to just do what I always do.* [blogger 4]

They Have Their Own Agenda

In addition to experiencing the mental health professionals as unsupportive, the bloggers (5/8) wrote about the professionals' lack of regard for the young adult's opinion. For example, "The appointment was fairly useless, apart from making me feel really crappy...The problem is that X knows very little about my current issues, purely because I didn't get to air them" (blogger 6).

Previous Lack of Support Contributed to My Problem

A few of the bloggers (3/8) identified that lack of support from others was a contributor to their current problems. For example:

> *I think the enormous lack of social contact and life experience is as much to blame for my low mood [as]...any physiological cause of depression.* [blogger 2]
>
> *I never got praised, I was never allowed to play with messy things, and if I cried I got told I was making a fool of myself, so that plus bullying throughout high school has [led] me to this point.* [blogger 6]
>
> *I think that the years of bad experiences I had at school has etched this "fear" into me, where I perhaps have an underlying issue that causes this "mental block" or disability.* [blogger 8]

I Have Moments of Connection

Young adult bloggers described feeling moments of connection as a result of their own initiative and the initiative of others.

Moments of Connection Because of Me

All bloggers (8/8) reported that when they initiated a conversation with friends, family, or a mental health professional they felt less lonely, which resulted in an improvement in their mood. The following are sample properties from the bloggers describing the positive impact they experienced through initiating connections with others:

> *Thing are beginning to look much better...I met our friends...all in all it was wildly entertaining, and I got to socialize with my boss [and others] so it was all rather relaxed and thoroughly enjoyable. Very beneficial to my mood state.* [blogger 3]
>
> *I've been feeling extremely hopeless and suicidal again. It often seems to disappear as soon as I get the chance to talk to someone about it and I feel like I'm being overly dramatic by bringing it up but this time I was determined to speak the truth about it.* [blogger 2]

Moments of Connection Because of Others

The majority of bloggers (7/8) also wrote about the positive impact of others reaching out to them. For example, "I felt very accepted and cared about by my family which is something I am unfamiliar with" (blogger 4).

Moments of Connection Because of Blogging

Blogging also provided a space in which young adults could connect with others in a safe and supportive environment.

Blogging Connects Me to Others Who Understand

Many of the bloggers (6/8) spoke about the supportive nature of blogging. That is, blogging allowed for the formation of interpersonal relationships and a sense of community, which often did not exist in the nonvirtual world. For example, blogger 7 wrote about the connections made through blogging (Textbox 3)

Textbox 3. Blogger 7's Description of Connections Made Through Blogging

> *For months I've been blogging about it; not thinking much about it, I've ended up knowing a lot of people who suffer of mental disorders. In the circle of depressive blogs, you find people who understand what you're going through. To be honest, without the Internet this would be rather difficult, if not impossible. When I see myself going through the worst, I think "what a pathetic, weak, (insert several other horrendous adjectives) person" despite my clinical knowledge of it. But when I read other bloggers going through the same, I want to hug them and for a second I see my own depressed self as someone worthy of the same support. You could say other blogs act as a mirror that is not being distorted by my own self-judgment.*

Blogging Allows Me to Help Others

Subsequently, some of the bloggers (3/8) felt that their writing and detailed accounts of mental health problems serve as a helpful resource tool for other people with mental health problems. For example:

> *One of the reasons I keep this blog is to give a different impression of what someone with a severe mental illness is like, to show that people like me are just ordinary people with mental illnesses and individual personalities.* [blogger 1]
>
> *Hopefully, whatever I leave on this blog will serve as a cautionary tale to anyone going through the same experiences and hopefully they will change before it's too late.* [blogger 2]

Discussion

The present study used qualitative methods to analyze unsolicited blog entries of young adults to learn about their experiences of living with a mental health problem and to explore their attitudes and beliefs about mental health and treatment. In summary, the young adults described very significant suffering and impairment resulting from their mental health problems. Their experience of living with a mental health problem can be summarized into two core categories: (1) I am powerless, and (2) I am utterly alone.

In terms of the core category *I am powerless* or without agency, young adults experienced their mental health problem as oppressive and overwhelming, yet they criticized themselves for not being able to control their difficulties and felt that they ought to be able to cope better. When seeking supports they experienced the mental health system as disempowering and controlling, and yet when they were able to engage in self-care they reported feeling better. In terms of isolation, young adults reported not connecting with others and that others did not connect with them. They noted feelings of profound isolation arising from shame, fear, guilt, and lack of ability to connect with others. Further compounding this isolation was that bloggers experienced others as failing to understand and failing to appreciate or show compassion for the difficulties the bloggers faced. Moreover, mental health professionals were experienced as unresponsive and unapproachable. Young adults perceived this lack of support and connection as contributing to their problems. Moments of connection, when they did occur, were experienced as restorative and as a source of hope.

Taken as a whole, the present results are consistent with the findings of previous research on young adults' mental health attitudes [35,36] and mental heath literacy [37,38]. For example, previous studies found that young adults, in contrast to other age groups, do not believe that it will be helpful to seek care for their mental health concerns [39]. This may be exacerbated when young adults have past negative experiences of seeking professional help, as described by some of the bloggers in the present study [35]. Moreover, previous research has shown that young adults are significantly less likely to report interest in receiving professional care, such as with primary care doctors and medications [36,40,41]. Rather, young adults prefer to handle their concerns on their own or with the support of friends and family [9,37,38,42].

Importantly, the present findings expand what we know from earlier research. For example, the profound feelings of powerlessness, struggle, loneliness, and isolation that these young adult bloggers write about has not been highlighted or deduced from existing quantitative studies [9,35]. The present study uncovered the experiential sense of living with a significant mental health problem—not just the young adults' attitudes, beliefs, or preferences [35,43,44]. In particular, while loneliness has been explored in past research [45,46], the profound feeling of aloneness described by the bloggers has not been highlighted in previous studies and perhaps should be a key goal of interventions or approaches to helping young adults with mental health problems. Mental health professionals, family, friends, and other allies should approach young adults with a firm sense of validation, understanding, empathy, and compassion.

Previous research has consistently found and documented that young adults prefer to handle mental health problems on their own and eschew the mental health system [37,38,47]. However, this previous work has not highlighted or investigated why this might be. Rather, the focus has typically been on developing ways to educate young adults about mental health treatment and convince them of the need for care and the benefits of care [48,49] in the absence of understanding why they might prefer to go it alone.

The present study suggests that one reason might be that young adults view the mental health system negatively and, for some, their experience has been consistent with this view. These findings strongly imply that the mental health care system should invest in efforts to educate others (lay people and mental health professionals alike) to create more welcoming, supportive environments that also facilitate choice in care. For example, not everyone wants medication [43,47,50] and it is imperative that there be engaged choice in treatment options, as this leads to higher treatment adherence and improved outcomes [51,52]. Efforts should also focus on creating and developing resources that allow young adults to feel connected (eg, blogging and informal supports). The creation of more accepting environments and attitudes will in turn facilitate greater self-acceptance among young adults, as an antidote to substantive self-criticism and self-blame. Indeed, the creation of more validating, empowering, and socially integrating mental health care treatments may be a more effective way to increase help seeking among young adults, compared with efforts aimed at improving mental health literacy.

Results of the present analysis also suggest that the act of blogging had several potentially therapeutic outcomes for the bloggers. First, the bloggers indicated that writing blogs was a way of expressing their inner emotions and difficulties. In this way writing provided a vehicle for the bloggers to reflect on their experiences in written form and gain understanding and sense of mastery over their problems. This finding is consistent with past research exploring the potential benefits of writing [53]. Many studies have shown that therapeutic writing can actually decrease symptoms of mental health concerns [53,54]

and improve physical health [55]. In fact, therapeutic writing is sometimes used in psychotherapy contexts to assist in improving insight and mental health [56-58]. Second, the bloggers spoke about how blogging was an important communication medium for them, especially when they experienced such a profound sense of disconnection from the rest of their lives. In this way, blogging can be useful for self-expression, sharing, and decreasing a sense of loneliness [59,60].

These two key reasons for blogging, therapeutic writing and social connections, are also reflected in Technorati's state of the blogosphere [61] study. This study of blogging more generally identified three distinct reasons for blogging: (1) self-expression, (2) sharing expertise and experiences, and (3) making money or doing business. The first two reasons are the most prominent motivators for blogging. Similarly, Nardi and colleagues [62] suggested five key motivations for blogging: (1) to chronicle life in order to share with others, (2) to express opinions and commentary in order to influence others, (3) to seek feedback and the views of others (eg, form participation), (4) to clarify or articulate one's own thinking through the act of writing, and (5) to express deep emotions and release tension. The benefits of using this vehicle, blogging, for self-expression and connections is worthy of further exploration in future studies. These factors could be particularly important to people, such as the young adult bloggers in the present study, who feel isolated and disconnected in their offline lives.

Limitations and Future Directions

The present study had a few important limitations. In particular, not all young adults with mental health problems choose to write blogs about their experiences. Further, the present sample was likely representative of higher symptom severity, and the bloggers were more likely treatment refractory, than the general population of young adult with mental health problems. As such, the participants were not representative of all young adults with mental health concerns. Future studies should be conducted with a broader range of young adults such as those who do not blog about their experiences and who have less severe problems or who may be at earlier stages of problem development or treatment.

The young adults examined in the present study have previously been difficult to hear from, and thus their voices have not shaped mental health treatment policies and programs. Moreover, the in-depth exploration of this particular group of young adults has generated important ideas for reforming mental health services for all young adults—namely, the creation of more validating, empowering, and socially integrating mental health care treatments. We suggest that future research ought to systematically examine what it is that young adults want to know about mental health problems and what treatment environments would be most conducive to young adults' seeking care. In our view such research programs would contribute greatly to solving the problem of poor utilization of mental health care services among young adults.

Conclusions

Even though young adults are very likely to have mental health difficulties, they are highly unlikely to seek or obtain mental health services, compared with other age demographics [63,64]. Thus, understanding the experiences of young adults with mental health problems, and their experiences of and attitudes toward receiving care, is critical to informing interventions and outreach efforts to better address these problems. The present study is one attempt to more fully understand these experiences, and the findings have several implications for meeting the needs of young adults with mental health problems. Through this infodemiology study, analysis of the bloggers' accounts can be used to inform improvements to public health for young adults experiencing mental health concerns [65]. The findings of the present study also support the value of Internet research for gaining insight and understanding into the lives of individuals with mental health illnesses [66]. Individuals with severe mental health problems are a very important population, as they are among the most distressed and costly to the mental health care system. As such, it is recommended that researchers continue to use Internet communication vehicles to reach out to this underserved and undertreated population for improving receptivity and knowledge of mental health care services, as well as building trust and awareness among the community at large for understanding the complexity of mental health illnesses.

Acknowledgments

Members of the Mobilizing Mind Research Group (mobilizingminds.ca) include the following (in alphabetical order). *Young adult partners:* Chris Amini, Meagan DeJong, Pauline Fogarty, Mark Leonhart, Kristin Reynolds, Allan Sielski, Tara Syed, and Alex Yaeger; *community partners:* Maria Luisa Contursi and Christine Garinger from mindyourmind (mindyourmind.ca); and *research partners:* Lynne Angus, Chuck Cunningham, John D Eastwood, Jack Ferrari, Patricia Furer, Madalyn Marcus, Jennifer McPhee, David Phipps, Linda Rose-Krasnor, Kim Ryan-Nicholls, Richard Swinson, John Walker, and Henny Westra.

The current study was supported by the Mobilizing Minds Project, a research team funded by the Canadian Institute of Health Research (grant number 180193). A special thanks to the bloggers for sharing their experiences.

Conflicts of Interest

None declared.

References

1. Statistics Canada. 2003. Canadian Community Health Survey: Mental Health and Well-being URL: http://www.statcan.gc.ca/daily-quotidien/030903/dq030903a-eng.htm [accessed 2011-05-19] [WebCite Cache ID 5ynZnIKh6]
2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602 [FREE Full text] [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
3. Wittchen HU, Nelson CB, Lachner G. Prevalence of mental disorders and psychosocial impairments in adolescents and young adults. Psychol Med 1998 Jan;28(1):109-126. [Medline: 9483687]
4. 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]
5. Edlund MJ, Wang PS, Berglund PA, Katz SJ, Lin E, Kessler RC. Dropping out of mental health treatment: patterns and predictors among epidemiological survey respondents in the United States and Ontario. Am J Psychiatry 2002 May;159(5):845-851. [Medline: 11986140]
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]
7. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
8. Armitage CJ, Conner M. Efficacy of the Theory of Planned Behaviour: a meta-analytic review. Br J Soc Psychol 2001 Dec;40(Pt 4):471-499. [Medline: 11795063]
9. Jorm AF, Barney LJ, Christensen H, Highet NJ, Kelly CM, Kitchener BA. Research on mental health literacy: what we know and what we still need to know. Aust N Z J Psychiatry 2006 Jan;40(1):3-5. [doi: 10.1111/j.1440-1614.2006.01734.x] [Medline: 16403031]
10. Eysenbach G, Kohler C. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. AMIA Annu Symp Proc 2003:225-229. [Medline: 14728167]
11. 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 2012-01-16] [WebCite Cache ID 64joo8gaC]
12. Fox S, Rainie L. Vital decisions: how Internet users decide what information to trust when they or their loved ones are sick. 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 2012-01-16] [WebCite Cache ID 64jotpK0T]
13. Mazur E. Teen blogs as mines of adolescent data. Teach Psychol 2005;32(3):180-182.
14. Hevern VW. Threaded identity in cyberspace: weblogs & positioning in the dialogical self. Identity 2004;4(4):321-335. [doi: 10.1207/s1532706xid0404_2]
15. NM Incite. 2011. BlogPulse URL: http://www.blogpulse.com/ [accessed 2011-05-19] [WebCite Cache ID 5ynaWNXbr]
16. Verna P. eMarketer. 2010 Sep. The Blogosphere: Colliding with Social and Mainstream Media URL: http://www.emarketer.com/Report.aspx?code=emarketer_2000708 [accessed 2011-06-13] [WebCite Cache ID 5zPdIwv5X]
17. Rainie L, Purcell K, Smith A. The social side of the internet. Washington, DC: Pew Internet & American Life Project; 2011 Jan 18. URL: http://www.pewinternet.org/~/media//Files/Reports/2011/PIP_Social_Side_of_the_Internet.pdf [accessed 2012-01-19] [WebCite Cache ID 64oVtUyuD]
18. Zickuhr K. Generations 2010. Washington, DC: Pew Internet & American Life Project; 2010 Dec 16. URL: http://www.pewinternet.org/~/media//Files/Reports/2010/PIP_Generations_and_Tech10.pdf [accessed 2012-01-19] [WebCite Cache ID 64oVw3UBH]
19. Lenhart A, Fox S. Bloggers: a portrait of the Internet's new storytellers. Washington, DC: Pew Internet & American Life Project; 2006 Jul 19. URL: http://www.pewinternet.org/~/media//Files/Reports/2006/PIP%20Bloggers%20Report%20July%2019%202006.pdf.pdf [accessed 2012-01-16] [WebCite Cache ID 64jp4BfTH]
20. Rainie L. The state of blogging. Washington, DC: Pew Internet & American Life Project; 2006 Jan. URL: http://www.pewinternet.org/~/media//Files/Reports/2005/PIP_blogging_data.pdf.pdf [accessed 2012-01-19] [WebCite Cache ID 64oVyE6ZR]
21. Baker JR, Moore SM. Blogging as a social tool: a psychosocial examination of the effects of blogging. Cyberpsychol Behav 2008 Dec;11(6):747-749. [doi: 10.1089/cpb.2008.0053] [Medline: 19072151]
22. Baker D, Fortune S. Understanding self-harm and suicide websites: a qualitative interview study of young adult website users. Crisis 2008;29(3):118-122. [Medline: 18714907]
23. Meyer D. Online self-help: developing a student-focused website for depression. Couns Psychother Res 2007 Sep;7(3):151-156. [doi: 10.1080/14733140701566357]
24. Whitlock JL, Powers JL, Eckenrode J. The virtual cutting edge: the internet and adolescent self-injury. Dev Psychol 2006 May;42(3):407-417. [doi: 10.1037/0012-1649.42.3.407] [Medline: 16756433]
25. Rosenthal R. Experimenter Effects in Behavioral Research. New York, NY: Appleton-Century-Crofts; 1966.
26. Brownlow C, O'Dell L. Ethical Issues for qualitative research in on-line communities. Disabil Soc 2002;17(6):685-694. [doi: 10.1080/0968759022000010452]

27. Clarke J, van Amerom G. A comparison of blogs by depressed men and women. Issues Ment Health Nurs 2008 Mar;29(3):243-264. [doi: 10.1080/01612840701869403] [Medline: 18340611]
28. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: 11701577]
29. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, IL: Sociology Press; 1967.
30. Hill CE, Thompson BJ, Williams EN. A guide to conducting consensual qualitative research. Couns Psychol 1997 Oct;25:517-572. [doi: 10.1177/0011000097254001]
31. Glaser BG. Theoretical Sensitivity: Advances in the Methodology of Grounded Theory. Mill Valley, CA: Sociology Press; 1978.
32. Rennie DL, Phillips JR, Quartaro GK. Grounded theory: a promising approach to conceptualization in psychology? Can Psychol 1988 Apr;29:139-150. [doi: 10.1037/h0079765]
33. Lincoln YS, Guba EG. Naturalistic Inquiry. Beverly Hills, CA: Sage Publications; 1985.
34. Alvesson M, Sköldberg K. Reflexive Methodology: New Vistas for Qualitative Research. Thousand Oaks, CA: SAGE; 1999.
35. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people's help-seeking for mental health problems. Aust e-J Adv Ment Health 2005;4(3):1-34.
36. Wilson CJ, Deane FP, Ciarrochi J. Can hopelessness and adolescents' beliefs and attitudes about seeking help account for help negation? J Clin Psychol 2005 Dec;61(12):1525-1539. [doi: 10.1002/jclp.20206] [Medline: 16173086]
37. Marcus M, Westra HA, Mobilizing Minds Research Group. Mental health literacy in Canadian young adults: results of a national survey. Can J Commun Ment Health 2012.
38. Wright A, Harris MG, Wiggers JH, Jorm AF, Cotton SM, Harrigan SM, et al. Recognition of depression and psychosis by young Australians and their beliefs about treatment. Med J Aust 2005 Jul 4;183(1):18-23 [FREE Full text] [Medline: 15992332]
39. Gonzales JM, Alegria M, Prihoda TJ. How do attitudes toward mental health treatment vary by age, gender, and ethnicity in young adults? J Community Psychol 2005;33:611-629. [doi: 10.1002/jcop.20071]
40. Boyd C, Francis K, Aisbett D, Newnham K, Sewell J, Dawes G, et al. Australian rural adolescents' experiences of accessing psychological help for a mental health problem. Aust J Rural Health 2007 Jun;15(3):196-200. [doi: 10.1111/j.1440-1584.2007.00884.x] [Medline: 17542793]
41. Rickwood D. The effectiveness of seeking help for coping with psychological problems in late adolescence. J Youth Adolesc 1995;24(6):685-703. [doi: 10.1007/BF01536951]
42. Schonert-Reichl KA, Muller JR. Correlates of help-seeking in adolescence. J Youth Adolesc 1996;25:705-730. [doi: 10.1007/BF01537450]
43. Bradley KL, McGrath PJ, Brannen CL, Bagnell AL. Adolescents' attitudes and opinions about depression treatment. Community Ment Health J 2010 Jun;46(3):242-251. [doi: 10.1007/s10597-009-9224-5] [Medline: 19636707]
44. Jorm AF, Wright A. Influences on young people's stigmatising attitudes towards peers with mental disorders: national survey of young Australians and their parents. Br J Psychiatry 2008 Feb;192(2):144-149 [FREE Full text] [doi: 10.1192/bjp.bp.107.039404] [Medline: 18245033]
45. Buchholz ES, Catton R. Adolescents' perceptions of aloneness and loneliness. Adolescence 1999;34(133):203-213. [Medline: 10234378]
46. Heinrich LM, Gullone E. The clinical significance of loneliness: a literature review. Clin Psychol Rev 2006 Oct;26(6):695-718. [doi: 10.1016/j.cpr.2006.04.002] [Medline: 16952717]
47. Hickie IB, Luscombe GM, Davenport TA, Burns JM, Highet NJ. Perspectives of young people on depression: awareness, experiences, attitudes and treatment preferences. Early Interv Psychiatry 2007 Nov;1(4):333-339. [doi: 10.1111/j.1751-7893.2007.00042.x] [Medline: 21352121]
48. Morgan A, Jorm A. Awareness of beyondblue: the national depression initiative in Australian young people. Australas Psychiatry 2007 Aug;15(4):329-333. [doi: 10.1080/10398560701323976] [Medline: 17612888]
49. Wright A, McGorry PD, Harris MG, Jorm AF, Pennell K. Development and evaluation of a youth mental health community awareness campaign - The Compass Strategy. BMC Public Health 2006;6:215 [FREE Full text] [doi: 10.1186/1471-2458-6-215] [Medline: 16923195]
50. Feeny NC, Zoellner LA, Kahana SY. Providing a treatment rationale for PTSD: does what we say matter? Behav Res Ther 2009 Sep;47(9):752-760. [doi: 10.1016/j.brat.2009.06.007] [Medline: 19616197]
51. Raue PJ, Schulberg HC, Heo M, Klimstra S, Bruce ML. Patients' depression treatment preferences and initiation, adherence, and outcome: a randomized primary care study. Psychiatr Serv 2009 Mar;60(3):337-343. [doi: 10.1176/appi.ps.60.3.337] [Medline: 19252046]
52. Wilder CM, Elbogen EB, Moser LL, Swanson JW, Swartz MS. Medication preferences and adherence among individuals with severe mental illness and psychiatric advance directives. Psychiatr Serv 2010 Apr;61(4):380-385. [doi: 10.1176/appi.ps.61.4.380] [Medline: 20360277]

53. Pennebaker J. Writing about emotional experiences as a therapeutic process. In: Salovey P, Rothman AJ, editors. Salovey P, Rothman AJ, editors. Social Psychology of Health: Key Readings in Social Psychology. New York, NY: Psychology Press; 2003.
54. Johnston O, Startup H, Lavender A, Godfrey E, Schmidt U. Therapeutic writing as an intervention for symptoms of bulimia nervosa: effects and mechanism of change. Int J Eat Disord 2010 Jul;43(5):405-419. [doi: 10.1002/eat.20714] [Medline: 19544556]
55. Mosher CE, Danoff-Burg S. Health effects of expressive letter writing. J Soc Clin Psychol 2006;25(10):1122-1139.
56. Goldberg D. Emplotment: letter writing with troubled adolescents and their families. Clin Child Psychol Psychiatry 2000;5(1):63-76. [doi: 10.1177/1359104500005001007]
57. Kerner EA, Fitzpatrick MR. Integrating writing into psychotherapy practice: a matrix of change processes and structural dimensions. Psychotherapy (Chic) 2007 Sep;44(3):333-346. [doi: 10.1037/0033-3204.44.3.333] [Medline: 22122259]
58. White VE, Murray MA. Passing notes: the use of therapeutic letter writing in counseling adolescents. J Ment Health Couns 2002;24(2):166-176.
59. Baker JR, Moore SM. Distress, coping, and blogging: comparing new myspace users by their intention to blog. Cyberpsychol Behav 2008 Feb;11(1):81-85. [doi: 10.1089/cpb.2007.9930] [Medline: 18275317]
60. Clarke BH. Early adolescents' use of social networking sites to maintain friendship and explore identity: implications for policy. Policy Internet 2009;1(1):55-89. [doi: 10.2202/1944-2866.1018]
61. Technorati. 2009. State of the Blogosphere 2009 URL: http://technorati.com/blogging/feature/state-of-the-blogosphere-2009 [accessed 2011-05-19] [WebCite Cache ID 5ynaJJTgY]
62. Nardi BA, Schiano DJ, Gumbrecht M, Swartz L. Why we blog. Commun ACM 2004;47(12):41-46. [doi: 10.1145/1035134.1035163]
63. Bebbington PE, Meltzer H, Brugha TS, Farrell M, Jenkins R, Ceresa C, et al. Unequal access and unmet need: neurotic disorders and the use of primary care services. Psychol Med 2000 Nov;30(6):1359-1367. [Medline: 11097076]
64. 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] [Medline: 18584997]
65. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
66. Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depress Anxiety 2011 Jun;28(6):447-455. [doi: 10.1002/da.20805] [Medline: 21400639]

Edited by G Eysenbach; submitted 13.06.11; peer-reviewed by K Boydell; comments to author 26.09.11; revised version received 08.11.11; accepted 17.11.11; published 30.01.12.

Please cite as:
Marcus MA, Westra HA, Eastwood JD, Barnes KL, Mobilizing Minds Research Group
What Are Young Adults Saying About Mental Health? An Analysis of Internet Blogs
J Med Internet Res 2012;14(1):e17
URL: http://www.jmir.org/2012/1/e17/
doi:10.2196/jmir.1868
PMID:22569642

Original Paper

Web-Based Recruiting for Health Research Using a Social Networking Site: An Exploratory Study

Yeshe Fenner[1,2], PhD; Suzanne M Garland[1,2,3,4], MBBS, MD, FRCPA, FAChSHM, FRANZCOG; Elya E Moore[2], PhD; Yasmin Jayasinghe[1,3,5], MBBS, FRANZCOG; Ashley Fletcher[1,2], MEpi; Sepehr N Tabrizi[1,2,3], PhD; Bharathy Gunasekaran[6]; John D Wark[7,8], PhD, MBBS, FRACP

[1]Department of Microbiology and Infectious Diseases, Royal Women's Hospital, Parkville, Australia
[2]Infection and Immunity Theme, Murdoch Childrens Research Institute, Parkville, Australia
[3]Department of Obstetrics and Gynaecology, University of Melbourne, Parkville, Australia
[4]Department of Microbiology, Royal Children's Hospital, Parkville, Australia
[5]Department of Gynaecology, Royal Children's Hospital, Parkville, Australia
[6]Melbourne Medical School, University of Melbourne, Parkville, Australia
[7]Department of Medicine, Royal Melbourne Hospital, University of Melbourne, Parkville, Australia
[8]Bone and Mineral Service, Royal Melbourne Hospital, Parkville, Australia

Corresponding Author:
John D Wark, PhD, MBBS, FRACP
Department of Medicine
Royal Melbourne Hospital
University of Melbourne
Department of Medicine (Royal Melbourne Hospital)
University of Melbourne
Parkville, 3050
Australia
Phone: 61 393427109
Fax: 61 393480796
Email: jdwark@unimelb.edu.au

Abstract

Background: Recruitment of young people for health research by traditional methods has become more expensive and challenging over recent decades. The Internet presents an opportunity for innovative recruitment modalities.

Objective: To assess the feasibility of recruiting young females using targeted advertising on the social networking site Facebook.

Methods: We placed an advertisement on Facebook from May to September 2010, inviting 16- to 25-year-old females from Victoria, Australia, to participate in a health study. Those who clicked on the advertisement were redirected to the study website and were able to express interest by submitting their contact details online. They were contacted by a researcher who assessed eligibility and invited them to complete a health-related survey, which they could do confidentially and securely either at the study site or remotely online.

Results: A total of 551 females responded to the advertisement, of whom 426 agreed to participate, with 278 completing the survey (139 at the study site and 139 remotely). Respondents' age distribution was representative of the target population, while 18- to 25-year-olds were more likely to be enrolled in the study and complete the survey than 16- to 17-year-olds (prevalence ratio = 1.37, 95% confidence interval 1.05–1.78, P = .02). The broad geographic distribution (major city, inner regional, and outer regional/remote) and socioeconomic profile of participants matched the target population. Predictors of participation were older age, higher education level, and higher body mass index. Average cost in advertising fees per compliant participant was US $20, making this highly cost effective.

Conclusions: Results demonstrate the potential of using modern information and communication technologies to engage young women in health research and penetrate into nonurban communities. The success of this method has implications for future medical and population research in this and other demographics.

(J Med Internet Res 2012;14(1):e20) doi:10.2196/jmir.1978

KEYWORDS

Advertising; research subject recruitment; women's health; Facebook

Introduction

Recruiting participants into health studies has become increasingly challenging. Traditional strategies, such as school-based recruitment, random digit dialing, systematic door knocking, and media advertising campaigns, have limitations including low participation rates [1], decreasing frequency of fixed household telephone line connections [2], and high costs [3,4]. Young people in particular are underrepresented in medical and population-based studies, as they are highly mobile, and recruitment and retention are difficult [5]. Modern social and technological changes have implications for research involving young people. A recent survey reported that 93% of 12- to 17-year-olds and 89% of 18- to 24-year-olds in the United States had access to the Internet [6], and the majority of these young people used the Internet daily [7]. An even more recent phenomenon is the dramatic rise in popularity of online social networking sites. The most popular social networking site is Facebook, with an estimated 800 million active users worldwide, of whom 50% will log on to Facebook in any given day [8]. While Facebook is already well established in developed nations, it is truly a global phenomenon, with the biggest growth in usage occurring in developing countries [9]. In Australia, use of social networking sites is the number one online activity for 16- to 29-year-olds, with 83% using them on a regular basis [10] and 93% of social networking site users being Facebook members [11]. Because social interactions between young people commonly occur via the Internet, social networking sites offer a promising new way to recruit participants, particularly young people, into medical research.

Web-based recruitment methods have been reported previously, including paid advertising and links on websites and online discussion boards [12-14]. We are aware of only a few publications describing health studies that used paid Facebook advertising to recruit participants [15-22]. Most of these studies grouped Facebook advertising with other online free and paid advertising strategies and did not compare the demographics of an exclusively Facebook-recruited sample with the target population. For instance, one study compared three methods for inviting young adult smokers to complete a survey [20]: (1) advertisements on the free online classifieds page Craigslist.org, (2) other Internet advertisements (including Facebook, MySpace, other social networking sites, Google, and file sharing and entertainment streaming websites), and (3) invitations to members of Internet market research panels. Method 2, which attracted younger participants and more males than the other methods, yielded the most completed surveys overall, while methods 1 and 3 were more cost effective and attracted participants more likely to complete the survey. However, the authors did not report demographic characteristics by Internet advertisement type. Another study that used Facebook advertisements in 2005 to invite US college students to complete a survey about prescription opioid misuse found that males and white students were more likely to respond to the advertisement; however, this may be consistent with the demographic profile of people who misuse prescription opioids [19]. At the time of that study, Facebook was far less popular than today with an audience of about 2.5 million users and open only to students with an educational email address (ie, extension .edu), making it difficult to generalize their results. Other studies have recruited participants by creating Facebook group pages and employing chain referral, or snowball sampling, to exploit group and friendship connections between Facebook users to obtain a convenience sample [15,23-25]. While this technique may be efficient and cost effective, it has limited potential to attract a representative sample, due to the reliance on social connections.

Our study differs from these previous studies in the following key ways: (1) recruitment at a time when the vast majority of the target population are regular Facebook users, thus giving the sampling modality a potentially broad reach, (2) systematic monitoring of each stage of recruitment including the display of, and response to, the advertisement, and navigation through our website (and as a function of age and regional group), and (3) assessment of the representativeness of a sample recruited exclusively using targeted Facebook advertising.

Our objectives were to assess (1) the feasibility of obtaining a representative sample of young females, using the Facebook targeted advertising system, which presents advertisements to users based on a selection of prespecified characteristics including location, age, and gender, and (2) young females' knowledge of and attitudes toward health issues, and participation in health and medical research. This was an exploratory study and we had no a priori hypotheses regarding these objectives.

Methods

Study Design

Inclusion criteria for participation in this cross-sectional study were (1) female, (2) 16–25 years old, (3) living in the Australian state of Victoria, and (4) willing to complete a health survey. The survey asked questions about demographic data, sexual and reproductive health, and willingness to participate in a larger health study. Exclusion criteria were perceived inability to give informed consent or complete the questionnaire due to inadequate understanding of the purpose and procedures of the study. We selected a target sample size of 200 as a reasonable number of participants to enroll within our budget and time frame.

Procedures

Facebook advertisements were displayed to Facebook users whose profiles matched our inclusion criteria: (1) between the ages of 16 and 25 years, inclusive, (2) female, and (3) located in Victoria, Australia. Age and gender are based on the information listed in the user's Facebook profile (age and gender are required by Facebook for all personal accounts), while location is based on the Internet protocol address or the address listed on the user's profile [26]. At the time of creating these Facebook advertisements (April 2010), city-level, but not

state-level, targeting was available for Australia. Therefore, we used city-level and geographic radius targeting [26] to target advertisements to people located within 50-mile radii of 17 cities throughout the state of Victoria. We used the largest possible radii and selected cities to maximize coverage of the state of Victoria. The Facebook advertisements comprised (1) one of several short titles (eg, "It's all about you," or "Tell us what you think"), (2) an image (eg, photos of young women of various ethnic backgrounds engaging in exercise or social activities), and (3) main text up to 135 characters in length (eg, "Are you 16–25 years and live in Victoria? We want to know what you think about health. Fill in a survey and go in a draw to win prizes," or "Tell us what health issues are important to you, fill in a survey and help improve the health and wellbeing of young Victorian women") (Figure 1).

Facebook gives advertisers the choice of being charged each time the advertisement is clicked (cost-per-click) or each time the advertisement is displayed a certain number of times (cost-per-thousand-impressions) [27]. We chose the cost-per-click option, as we were interested in people clicking through to our website. The advertiser also chooses a bid, which is the maximum the advertiser will pay for each click on the advertisement, in the cost-per-click model. From the available ad inventory, the Facebook advertising algorithm automatically selects the best advertisement to run based on advertisement performance and the cost-per-click or cost-per-thousand-impressions. Facebook advertisements compete with each other to appear in the ad space on the right-hand side of the webpage. For each advertisement, Facebook gives a suggested bid range, which is the range of bids currently winning the auction among similar advertisements being displayed to the targeted audience; a low bid makes it unlikely that the advertisement will be displayed [27]. Our bids ranged from US $0.70 to US $1.15 and fell within Facebook's suggested bid range. Our daily budget (the maximum Facebook charge per day; once reached, Facebook stops running the advertisements for that day) ranged from US $20 in the first week of advertising up to US $90 in the final week of advertising, and was adjusted according to our desired recruitment rate.

Advertisements appeared on Facebook from May 19 to September 29, 2010. From May 19 to June 29, 2010, we conducted a single advertising campaign, targeting all female Victorian Facebook users aged 16–25 years. Subsequently, we used six separate advertising campaigns to target each combination of three age groups (16–17, 18–21, and 22–25 years) and two regions (urban and nonurban), to obtain more detailed demographic information and allow the advertising budget for each campaign to be adjusted, if needed, to try to yield a representative sample. For instance, if fewer advertisers were competing to advertise to female adolescents under 18 years of age, having one single advertising campaign for 16- to 25-year-olds may lead to our advertisements being displayed preferentially to under 18s, due to the nature of Facebook's ad bidding process, without any way for us to control this. As above, we used city-level and geographic radius targeting parameters to target advertisements to people in urban and nonurban Victoria. For the urban advertising campaigns, advertisements were targeted at females located within 10 miles of Melbourne, the capital city of Victoria. For the nonurban advertising campaigns, advertisements were targeted at females located within 50 miles of 12 cities, an area that covered most of regional Victoria while excluding major cities.

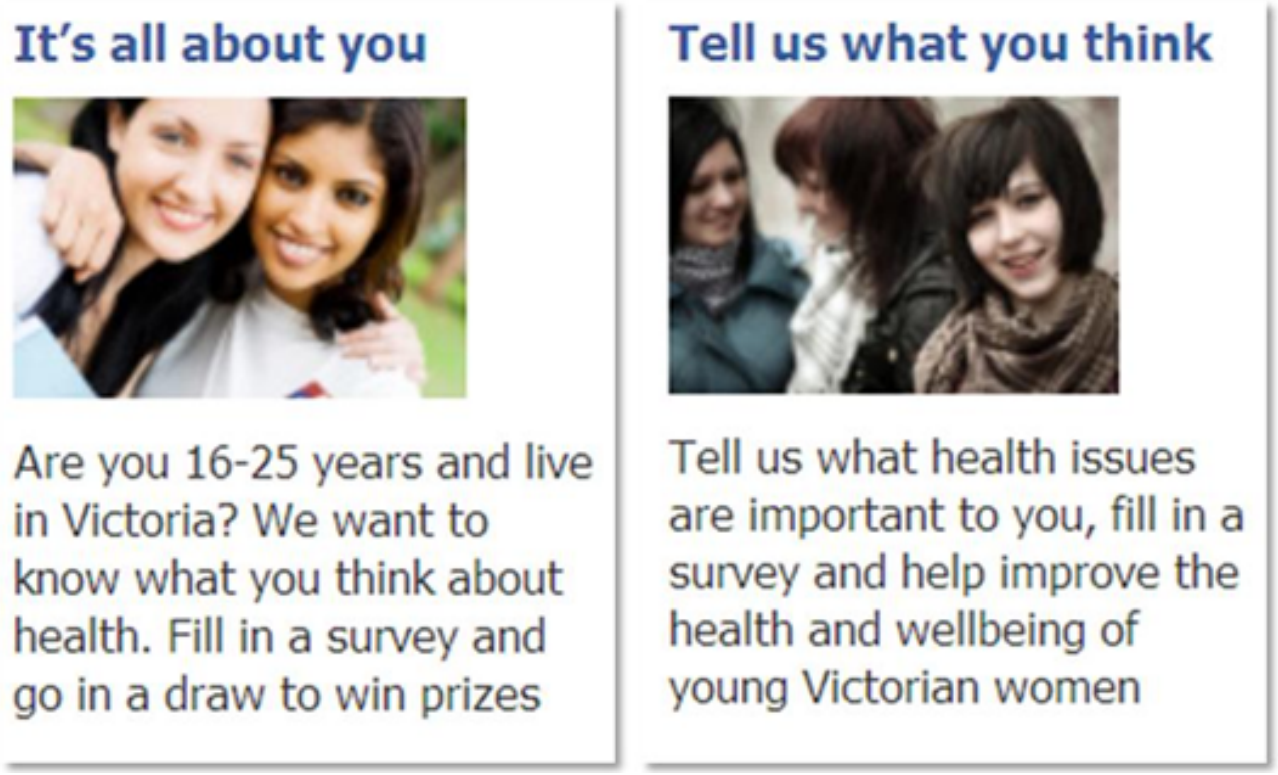
When a Facebook user clicked on an advertisement, she was redirected to our secure study website (www.yfhi.org), containing study information (details about the objectives and procedures of the study, eligibility criteria, prize draw for completing the survey, and researchers' backgrounds, affiliations, and contact information) and an expression-of-interest form to enable users to learn more about the study.

Potential participants who visited our study website could either send us their telephone and email contact information through a secured encrypted online system, or directly contact study personnel. Research staff assessed eligibility of all participants over the telephone after initial contact, explained the study, and assessed respondents' competence to give informed consent (based on their ability to understand the purpose and procedures of the study and explain it in words). Those eligible and interested were invited to visit a study site in the suburb of North Melbourne, Victoria, to complete a health-related survey. The study site was an office suite located 2 km north of the city center of Melbourne, Victoria, in a medical precinct with several major hospitals situated nearby and readily accessible by public transport. Respondents who declined to visit the study site were then invited to complete the survey online remotely. The reason for initially inviting respondents to visit our study site was to assess the proportion of young females who would travel to participate, which informs the suitability of this sampling method for recruiting females into future studies requiring in-person contact. Respondents under 18 years of age underwent a mature minor assessment by a researcher following guidelines from the Medical Practitioners Board of Victoria, *Consent for Treatment and Confidentiality in Young People* [28], on how to define a mature and competent young person. Briefly, the researcher assessed age, general maturity of speech, level of schooling, and ability to understand the nature and rationale of the research project and to explain it in words. Verbal consent was obtained from all participants by telephone. In addition, written consent was obtained from those participants visiting the study site.

Respondents were considered unreachable after no response to three missed telephone calls, plus an SMS message and/or email. Respondents who initially consented to participate were considered lost to follow-up if they were unreachable to schedule an appointment at the study site, or did not complete the online survey remotely after three reminder emails. Respondents who were initially consented into the study site group, but who subsequently were unable or unwilling to visit the study site, were given the option of completing the survey remotely.

Participants were offered AU $25 (AU $1 = US $1.05 in April 2011) compensation for their time and travel costs if they visited the study site (and up to AU $70 additional travel reimbursement if travelling from regional areas) or AU $15 compensation for their time for completing the survey remotely.

Figure 1. Examples of Facebook advertisements.



Participants who visited the study site were asked to complete the online survey at a computer in a private room. Those who participated remotely were emailed instructions for accessing the online survey. We used the online survey tool Survey Monkey (www.surveymonkey.com), with the enhanced security option of secure socket layers encryption, to administer the survey to participants. To further protect privacy, we masked participants' Internet protocol addresses so that they were not stored in the survey results. To enable compliance monitoring, researchers provided participants with a unique study identification number to access the survey. The survey contained questions about demographic variables (date of birth, marital status, living arrangements, income, country of birth, education, employment status, indigenous status, ethnicity, and postal/zip code), height and weight, how they found out about the study, sexual history, experience and knowledge of sexually transmitted infections (*Chlamydia trachomatis* and human papillomavirus), and the acceptability of participating in a long-term research study, including answering sensitive questions and undergoing physical examinations (Multimedia Appendix 1).

Statistical Analysis

Statistical analyses were performed using Stata version 11.1 (StatCorp LP, College Station, TX, USA). We used Australian Bureau of Statistics 2006 census data [29] and Victorian Population Health Survey 2008 data [30] to compare our cohort with the general population. Socioeconomic status was assigned using the Bureau's Socio-Economic Indexes For Areas (SEIFA) and the 2006 Postal Area Index of Relative Socio-economic Advantage and Disadvantage, which is a continuum of advantage (high values) to disadvantage (low values) scores [31].

We compared sociodemographic characteristics (age group, geographic region, country of birth, indigenous status, socioeconomic level, and education level) and body mass index (BMI) in our sample with that of the general population using a Fisher exact test.

Prevalence ratios (PRs), 95% confidence intervals (CIs), and 2-sided *P* values were estimated using log-binomial regression [32]. When the log-binomial model failed to converge, we used a Poisson model with robust error variance as an approximation [32]. We estimated PRs of clicking on the advertisement, mutually adjusting for geographic region and age group, and PRs of submitting an expression of interest, stratifying by age group. Associations with visiting the study site to complete the survey, rather than completing it online remotely, were also estimated using PRs, mutually adjusting for age group, geographic region, country of birth, socioeconomic level, education level, and BMI.

In all analyses, we defined a 2-sided *P* value of <.05 as statistically significant. Data were treated as missing if no response was given or "don't know" was selected.

Ethical Considerations

We obtained ethical approval for the study through the Human Research and Ethics Committees at the Royal Women's Hospital, Melbourne, Australia, and adhered to the *National Statement on Ethical Conduct in Human Research* [33], which was developed to protect the interests of people who participate in research studies.

The confidentiality of all participants was maintained throughout the study. Unique codes for participant identification were used, data were stored in password-protected computers and files, and data transmitted electronically were securely encrypted.

Facebook uses an automatic advertising system, in which no individual user's information is revealed to the advertiser. After clicking on an advertisement, users of Facebook were automatically directed to our secure study website, and all subsequent study procedures took place outside Facebook. This approach minimized the amount of information exchanged via the Facebook website, to further ensure the privacy and security of participants' information.

Results

Recruitment

The Facebook advertisements were displayed 36,154,610 times, resulting in 8339 clicks on the advertisement (which directed respondents to the study website) and 551 expressions of interest submitted through our website. The number of times an advertisement was displayed to a unique Facebook user was 469,678, resulting in 7940 unique clicks (some Facebook users clicked on the advertisement multiple times, bringing the total number of clicks to 8339). In total 65.69% (3121/4751) of logged visits to our website lasted less than 10 seconds, 12.0% (568/4751) lasted between 10 seconds and 1 minute, and 22.35% (1062/4751) lasted for more than 1 minute. The About This Study webpage, containing details about the aims of the study and eligibility criteria, received 1144 unique visitors, who spent an average of 53 seconds on the page.

Of the 551 participants who initially responded to an advertisement, 426 were contactable by telephone and enrolled in the study (none was excluded due to not meeting the eligibility criteria), and 278 completed the survey (Figure 2), which took most participants 15–30 minutes to complete. Thus, the participation rate for those who clicked on the advertisement was 3.5% (278/7940), and the participation rate for those who read about the study from the About This Study webpage was 24.3% (278/1144). The average Facebook charge was US $0.67 per click, amounting to $10.16 per expression of interest, or $20.14 per compliant participant. Age and geographic region were not strong predictors of the likelihood of clicking on the advertisement. However, for those who did click on the advertisement, older age was predictive of submitting an expression of interest (Table 1). Because of this, as well as differences in average bids for each campaign, the average cost per participant varied with age group ($15, $23, and $49 per participant for 22- to 25-, 18- to 21-, and 16- to 17-year-olds, respectively, using data from June 30, 2010 onward, when there were separate advertising campaigns for the different age groups).

The age distribution of the 551 initial respondents who submitted an expression of interest reflected the general population (Table 2). This was achieved despite the lower odds of 16- to 17-year-olds submitting an expression of interest after clicking on the advertisement, by targeting the Facebook campaign budget to elicit more clicks from 16- to 17-year-olds. The broad geographic distribution, measured by Remoteness Area [29], of the 328 (59.5%) initial respondents who provided their location revealed moderate overrepresentation of regional/rural females.

Table 1. Prevalence ratios of clicking on the Facebook advertisement and submitting an expression of interest

Characteristic	Clicking on an advertisement			Expression of interest		
	Adjusted PR[a]	95% CI[b]	*P* value	PR[c]	95% CI	*P* value
Age group (years)						
16–17	1.00			1.00		
18–21	1.05	0.99–1.11	.10	2.42	1.85–3.19	<.001
22–25	1.14	1.07–1.21	<.001	3.34	2.56–4.36	<.001
Geographic region						
Major city	1.00					
Regional/rural	0.96	0.90–1.04	.36			

[a] Prevalence ratios (PRs) of clicking on a Facebook recruitment advertisement, mutually adjusted for geographic region and age group.

[b] Confidence interval.

[c] PRs of submitting an expression of interest, after clicking on an advertisement. Geographic region is omitted from this analysis because 40.5% (223/551) of respondents did not provide this information.

Figure 2. Summary of sampling and response. Percentages are calculated using the number of expressions of interest (551) as the denominator.

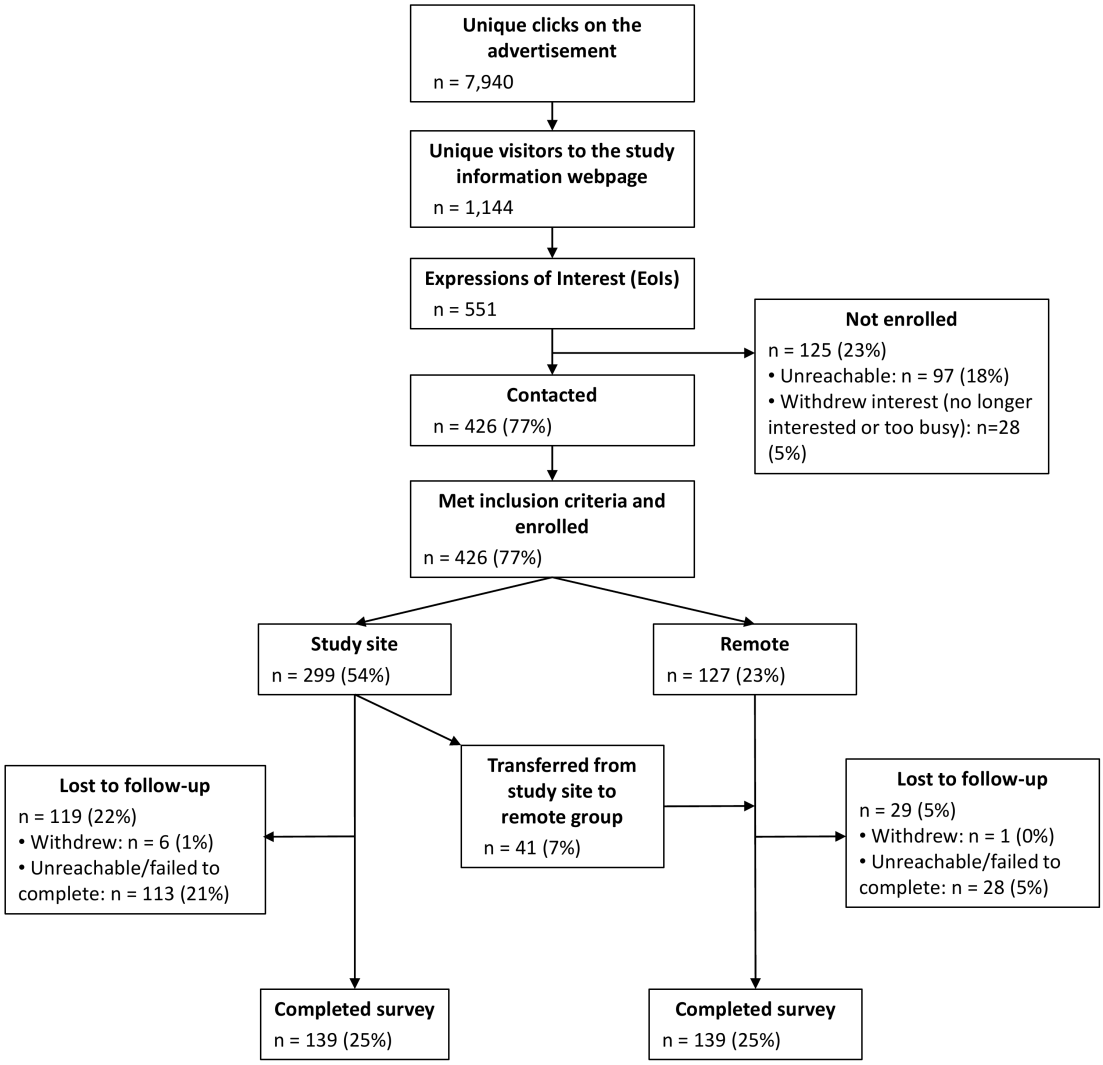


Table 2. Demographic characteristics of respondents who submitted an expression of interest

Characteristic	Respondents (n = 551)			Target population[b]	Fisher's exact *P* value
	n	%	95% CI[a]		
Age group (years)					
16–17	98	18%	14.6–21.0	19.8%	
18–21	217	39.4%	35.3–43.5	40.1%	
22–25	236	42.8%	38.7–47.0	40.1%	0.34
Geographic region [c]					
Major city	238	72.6%	67.7–77.4	78.7%	
Inner regional	72	22%	17.5–26.4	17.7%	
Outer regional/remote	18	6%	3.0–8.0	3.6%	0.02

[a] Confidence interval.

[b] Population data from Australian Bureau of Statistics census 2006, with figures corrected for nonresponses to add up to 100%.

[c] Only 328 of the 551 respondents provided geographic region information.

Participant Characteristics

Predictors of participation were older age, higher education level, and higher BMI (calculated from self-reported height and weight), as compared with the general target population (Table 3). Females born outside Australia and those from regional and lower socioeconomic areas were well represented, with these variables not associated with likelihood of participation.

Three participants identified themselves as Aboriginal or Torres Strait Islander Australians (1.08%, 95% CI 0.06–2.10), which is consistent with the 0.85% in the target population (*P* = .68).

Initial respondents 18–25 years old were more likely to be enrolled in the study and complete the survey than younger respondents (PR = 1.37, 95% CI 1.05–1.78, *P* = .02). This is associated with a higher proportion of 16- to 17-year-olds being unreachable, even using a combination of telephone calls, SMS, and email (29% (28/98) vs 15% (69/453), *P* = .003). Of the respondents who were contactable and enrolled, the completion rate did not vary by age group (*P* = .6).

The strongest predictor of willingness to travel to the study site to complete the survey, compared with completing it online from a remote location, was proximity to the study site, as measured by geographic region (Table 4). As a result, people from major cities were overrepresented in the study site population. However, the overall study population (study site plus remote) was geographically representative. Participants from areas with postal/zip codes in the highest bracket of socioeconomic advantage were 50% more likely to visit the study site than those in the lowest bracket, although there was no such difference for participants in the middle bracket (Table 4). This may be associated with proximity to the study site. Considering only the females living in the major cities region (in which our study site was located), the mean distance to the study site from participants' postal codes was 14, 29, and 39 km for those in the highest, middle, and lowest socioeconomic bracket, respectively. Furthermore, of the 211 participants in the major cities region, 78% (51/65; 95% CI 68%–88%) of those living within 10 km of the study site travelled to the study site to complete the survey, compared with only 51% (75/146; 95% CI 43%–59%) of those living further than 10 km from the study site. Indeed, when we included distance to the study site as a variable in our model, it was a significant predictor of attending the study site (PR = 0.92, 95% CI 0.84–0.99, *P* = .04, where each 10 km increase in distance from the study site corresponds to a factor 0.92 decrease in PR of visiting the study site, holding all other variables in the model constant), while geographic region and socioeconomic bracket were no longer significant.

Participants classified as overweight according to their BMI were less likely to travel to the study site than those with normal BMI, although participants who were obese and those with normal BMI were equally likely to visit the study site (Table 4). The prevalence of overweight/obesity (derived using the World Health Organization classifications of adult body weight status based on BMI [34]) increased with age, with 24% (9/37; 95% CI 10.5–38.1) of 16- to 17-year-olds, 33% (36/108; 95% CI 24.4–42.2) of 18- to 21-year-olds, and 36% (44/121; 95% CI 27.8–44.9) of 22- to 25-year-olds being overweight or obese. From a linear regression model, each added year in age corresponded to an increase in BMI of 0.29 kg/m^2 (95% CI 0.06–0.52, *R2* = .02, *P* = .01). Those from regional/rural areas were also more likely to be overweight or obese, with prevalences of 31% (63/203; 95% CI 24.7–37.4), 40% (18/45; 95% CI 25.7–54.3), and 50% (8/16; 95% CI 25.5–74.5) for major city, inner regional, and outer regional/remote areas, respectively.

The study questionnaire also asked nonurban participants to rate the likelihood that they would have a "physical examination and/or tests once per year for 4 years as part of a health study if" (1) "we paid for your travel to and from a study site in Melbourne and accommodation for up to two nights," and (2) "we travelled to your town to conduct the physical examination." These items were measured on a 5-point Likert-type scale, from 1 (not at all likely) to 5 (very likely). The mean rating was 4.49 (SD 0.84) for scenario 1 and 4.50 (SD 0.82) for scenario 2, indicating that these options were equally acceptable to nonurban participants.

Less than 5% of survey data were missing on the demographic variables presented, while less than 8% of data were missing on any single variable.

Table 3. Demographic characteristics of participants

Characteristic	Study population (n = 278)			Target population[c]	Fisher's exact *P* value
	n[a]	%	95% CI[b]		
Age group (years)					
16–17	38	14%	9.60–17.7	19.8%	
18–21	115	41.4%	35.5–47.2	40.1%	
22–25	125	44.9%	39.1–50.8	40.1%	.02
Geographic region					
Major city	211	76.5%	71.4–81.5	78.7%	
Inner regional	49	18%	13.2–22.3	17.7%	
Outer regional/remote	16	6%	3.0–8.6	3.6%	.15
Country of birth					
Australia	230	83.3%	78.9–87.8	80.2%	
Other	46	17%	12.2–21.1	19.8%	.23
Indigenous status					
Aboriginal or Torres Strait Islander	3	1%	0.06–2.10	0.85%	
Other	275	98.9%	97.9–99.9	99.15%	.51
Socioeconomic level (SEIFA percentile) [d]					
<55	83	30%	24.6–35.5	33.9%	
55–80	87	32%	26.0–37.0	32.7%	
>80	106	38.4%	32.6–44.2	33.4%	.19
Education level					
< Year 12[e]	54	19%	15.5–23.3	28.4%	
Year 12	104	37.4%	32.6–42.2	38.1%	
> Year 12	120	43.2%	37.3–49.0	33.5%	<.001
Body mass index (kg/m ^{2}), 18- to 24-year-olds [f]					
<18.5 (underweight)	13	6%	3.6–8.4	9.6%	
18.5–25 (normal)	128	61.0%	55.6–65.5	68.0%	
25–30 (overweight)	44	21%	16.6–24.8	15.8%	
>30 (obese)	25	12%	9.4–16.2	6.6%	.002

[a] Numbers may not add up to 278 due to missing data.

[b] Confidence interval.

[c] Population data from Australian Bureau of Statistics census 2006, except for body mass index data from the Victorian Population Health Survey 2008, with figures corrected for nonresponses to add up to 100%.

[d] Based on postal/zip code. Percentiles are the rankings within Victoria. Note that the percentiles are based on the postal codes and are not weighted by the population within each postal code. Major city postal codes have, on average, higher Socio-Economic Indexes for Areas (SEIFA) level and larger population than regional postal codes. Consequently, the population-weighted median SEIFA percentile is about 70%, not 50%.

[e] Year 12 is the final year of high school in the Australian education system.

[f] Age range chosen to match that from Victorian Population Health Survey 2008. Consistent with the Survey, we used the World Health Organization classifications of adult body weight status based on body mass index.

Table 4. Associations between completing[a] the survey at the study site and completing it remotely, by sociodemographic characteristics

Characteristic	Completed remotely (n = 139)[b]	Completed at study site (n = 139)[b]	Adjusted PR of visiting study site[c]	95% CI[d]	*P* value
Age group (years)					
16–17	26 (19%)	12 (8.6%)	1.00		
18–21	49 (35%)	66 (47.5%)	1.32	0.8–2.19	.28
22–25	64 (46%)	61 (43.9%)	1.15	0.68–1.95	.60
Geographic region					
Major city	85 (62%)	126 (90.7%)	1.00		
Inner regional	38 (28%)	11 (7.9%)	0.46	0.25–0.85	.01
Outer regional/remote	14 (10%)	2 (1.4%)	0.26	0.07–0.98	.05
Country of birth					
Australia	114 (83.2%)	116 (83.5%)	1.00		
Other	23 (17%)	23 (16.5%)	0.88	0.65–1.21	.44
Socioeconomic level (SEIFA percentile) [e]					
<55	59 (43%)	24 (17.3%)	1.00		
55–80	46 (34%)	41 (29.5%)	1.09	0.73–1.63	.68
>80	32 (23%)	74 (53.2%)	1.48	1.03–2.13	.03
Education level					
< Year 12[f]	37 (27%)	17 (12.2%)	1.00		
Year 12	41 (30%)	63 (45.3%)	1.34	0.86–2.10	.20
> Year 12	61 (44%)	59 (42.5%)	1.15	0.72–1.84	.55
Body mass index (kg/m ^{2})					
<18.5 (underweight)	12 (9%)	4 (3.0%)	0.49	0.22–1.10	.08
18.5–25 (normal)	70 (53%)	91 (67.9%)	1.00		
25–30 (overweight)	33 (25%)	22 (16.4%)	0.72	0.53–0.98	.04
>30 (obese)	17 (13%)	17 (12.7%)	1.05	0.74–1.50	.78

[a] For the purposes of this study, we define a survey as being complete if 80% of the demographic information needed for our analysis was provided. A total of 5 participants did not fully complete the survey, but did provide most of the demographic data used in our analyses.

[b] Numbers may not add up to 278 due to missing data.

[c] Prevalence ratios (PRs) of visiting study site to complete the survey, versus completing online remotely. Poisson regression models were mutually adjusted for age group, geographic region, country of birth, socioeconomic level (Socio-Economic Indexes for Areas [SIEFA] percentile), education level, and body mass index. Small numbers of indigenous females in our sample did not support meaningful analyses and thus indigenous status was excluded from this model.

[d] Confidence interval.

[e] Socio-Economic Indexes for Areas (SEIFA) percentile, based on postal/zip code.

[f] Year 12 is the final year of high school in the Australian education system.

Discussion

This study demonstrated good levels of engagement of young females, who are traditionally underrepresented in health studies or have poorer access to health care. For example, the strong representation of regional and rural females in this study shows the potential benefit of using social networking sites to recruit a segment that traditionally has been quite difficult to reach. Rural and regional participants were less likely to travel to the study site, which in most cases would have involved round trips of 2 to 8 hours. Nonetheless, a representative study site population could be obtained by oversampling nonurban females. Moreover, the survey results indicate that study site compliance rates may be increased by providing sites close to participants' place of residence or offering accommodation and compensation for regional participants to travel to an urban study site.

We obtained a representative distribution of Australian and non-Australian-born participants, which compares favorably with many population-based studies where overseas-born

participants are underrepresented. For instance, in the Australian Longitudinal Study on Women's Health (ALSWH), which used mail-out invitations in 1996 to recruit women randomly selected from the Medicare database, 88.6% of 18- to 23-year-old respondents were Australian-born versus 77.8% in the target population [35]. More recently, in the Victorian Population Health Survey 2008, which used random digit dialing to sample from residential households with landline telephone connections, 79.2% of respondents were Australian-born versus the target of 71.3% [30]. On the other hand, more highly educated females were overrepresented in our sample, which is another common bias in population-based studies (eg, the ALSWH, where 18- to 23-year-old respondents were more likely than the target population to be tertiary educated) [35]. The overrepresentation of more highly educated females in our study population may be overstated in the data. It is likely that some participants misunderstood the question about education level ("What is your highest level of education completed?") and indicated the level of education they are currently *completing*, rather than their highest level *completed*. For instance, of the 33 participants who indicated that they were currently attending high school, 8 (24%) answered that their highest level of education completed was Year 12. However, if that were the case, they would no longer be attending high school. We corrected these data, but we could not identify inconsistent answers for other educational levels, since someone may have, for example, completed a university degree and still be attending university.

Respondents 16–17 years old were less likely than those 18–25 years old to be enrolled in the study. The lower participation rate of 16- to 17-year-olds was associated with their being harder to contact, even using a combination of calling their mobile/cellular telephone and sending SMS messages and emails. A contributing factor may have been that the delay between the expression of interest and first attempt at contacting the 16- to 17-year-olds was, on average, 3 weeks longer than for 18- to 25-year-olds. This delay was due to the logistics of performing the mature minor assessment.

Overweight and obese young females were strongly represented. The average BMI (based on self-reported weight) in our study population was higher than that of the target population, with 33% (69/210) of 18- to 24-year-olds being overweight or obese, compared with 22% in the target population ($P < .001$). Previous reports suggest that Internet and interactive media use is positively correlated with BMI in adults [36] and adolescent females [37]. Our findings demonstrate the potential utility of Facebook as a recruitment tool in these high-risk young females, who may find initial online engagement less confrontational than other approaches. Notably, the prevalence of overweight and obesity combined rose appreciably across age groups, which is consistent with major lifestyle and health changes occurring at this stage of life. This observation also lends support to the urgent need for more research into evolving health risks in young women.

Social networking site recruitment shows great potential to yield a demographically representative sample by oversampling and appropriately weighting data. The ability to create multiple advertising campaigns targeted to different populations, and to closely monitor their real-time performance, allows one to reallocate resources between campaigns to direct recruitment efforts toward targeted parameters such as age and location of residence. Moreover, it seems clear from our findings that this approach could be used to direct recruitment campaigns to people with particular health problems or health risks.

At an average of US $20 in advertising fees per participant, this recruitment method compares favorably with traditional methods, which can cost US $20–$500 per participant, depending on the particular strategy and target population [3,4,38,39]. Traditional passive recruitment through paid media campaigns (eg, radio, television, and newspaper advertising) may have broad reach throughout the community but can be expensive [3] and ineffective, as proportional use of these media has decreased over the years compared with Internet use. Active face-to-face recruitment (eg, through schools, community groups, and health professionals) is generally labor intensive, expensive, and unsuitable for obtaining a representative population. Random digit dialing and direct mail that draws on information from electoral lists or health care databases have been widely used to obtain population-based samples, but as with other active recruitment methods, they are more labor intensive and costly than passive Facebook recruiting [40]. Random digit dialing telephone sampling has been a very popular recruitment and survey tool; however, its coverage is decreasing as fewer households have active landline telephones. In 2008, 16% of all US adults, 63% of adults in shared households (living with nonrelatives), and 31% of 18- to 24-year-olds lived in mobile/cellular phone-only households [2]. In Australia in 2010, 33% of all 18- to 24-year-olds had no landline telephone in their household, and for those living outside the parental home, the figure was almost 60% [41]. This trend is occurring worldwide, threatening the generalizability of studies that employ random digit dialing of landlines only. There have been promising results from random digit dialing studies that include mobile/cellular phones, but the costs were up to 5 times higher than for landline random digit dialing [42]. It should be noted that Facebook advertising costs are likely to increase, particularly as it grows in popularity, with more advertisers competing for ad space and driving up the bid price. Indeed, the cost-per-click rates in the United States, United Kingdom, France, and Germany rose 74% in the 12 months from the second quarter of 2010 to the second quarter 2011 [43].

Strengths and Limitations of the Study

A limitation of this recruitment method is the low participation rate and the potential for volunteer bias. The participation rate of those who visited our About This Study webpage and had an opportunity to read about the study and make an informed decision about joining was 24%, while the participation rate of those who clicked on the Facebook advertisement was only 3.5%. This is lower than typical rates for population-based studies using traditional recruitment methods. For instance, using mail-out recruitment, the ALSWH had a response rate of about 40% in its youngest cohort [35], which is typical for mail-out survey response rates [44]. ALSWH compared their sample demographics with census data to confirm that the participants were reasonably representative of the general population. Despite their higher response rate, they actually

found evidence for a bigger response bias than in our study, and in particular, a more serious overrepresentation of tertiary-educated females (12.1% had completed a university degree vs 7.7% in the target population) and Australian-born females (88.6% vs 77.8%). Broad demographics aside, there may be biases in our sample that we did not measure with our survey—for instance, psychological, social, or familial factors—although we did use the SEIFA as an approximation of socioeconomic advantage and found no difference between the study and target populations. Further studies are needed to investigate other predictors of participation, and information on reasons for nonparticipation would also be informative.

Other limitations include bias due to exposure to the advertisement being positively correlated with time spent on Facebook (and therefore not the same for each user), the need for users to supply their correct gender and age in their profile in order to be exposed to the advertisement, and chain sampling bias, whereby users exposed to the advertisement may share information about the study with others who may then submit an expression of interest. We were able to evaluate the latter phenomenon and found that it did not meaningfully change our results. Specifically, 25 participants indicated that they found out about the study through a friend or relative. Exclusion of these females from the analyses did not result in any significant change in the results. We did not omit them from the main analysis because in observational studies, particularly in this age group and in this era of rapid information sharing, we expect this to be a common occurrence, although it is not always measured in research studies.

This study was conducted in the state of Victoria, Australia, and further research is needed to determine how applicable our findings are to health research internationally. Several factors suggest that our results may be more broadly generalizable. In particular, most other developed nations are also witnessing the phenomenon of very high rates of social networking site usage along with declining landline telephone prevalence, which together make this a very promising recruitment approach that is readily portable to other regions and countries. Moreover, as above, social networking site use in developing countries is expanding at very high rates, suggesting the potential to use this approach to recruitment for biomedical research in many parts of the world.

As this was a cross-sectional study, we were unable to measure retention rates, which would have informed applications of this sampling frame to longitudinal studies. There was a high level of self-reported willingness to answer questionnaires and have physical examinations, tests, and sample collections in an ongoing health study. Indeed, nearly all respondents stated that they would like to be contacted about participating in a large longitudinal study about young women's health.

Despite the above limitations, the study population was demographically similar to the general population of 16- to 25-year-old Victorian females. We obtained a large sample size (n = 278) for the recruitment period and achieved high compliance in completing the survey with very little missing data, and the method was highly cost effective. In addition, this method allows researchers to very precisely extract the components of responses (clicks on the advertisement, navigation around the website, expression of interest, etc), whereas many other recruitment methods do not.

Conclusions

Results from this study suggest that targeted recruitment using the social networking site Facebook has strong potential for yielding demographically representative samples of young females, a population far more likely to engage with the health community using Facebook than landline telephones. A substantial majority of participants expressed clear willingness to participate in extensive longitudinal studies of their health. This model may also be appropriate for recruiting other populations, as use of social networking sites by older people and minorities continues to grow [45,46], as well as for intervention studies and clinical trials. Furthermore, this strategy was cost effective in comparison with many traditional methods.

Acknowledgments

The authors wish to thank Advanced Medical Science (AMS) student Navera Ahmed for her considerable help in conducting the study and disseminating results. We are grateful to the members of our Young Women's Advisory Group for their feedback about the study website, logo, advertisements, and questionnaire. We also thank all the young women who took part in this study.

This work was supported by a grant from the Dean Robert Winter Trust, which is managed by the ANZ Trustees, the National Health and Medical Research Council (Program Grant #568971), and the Victorian Government's Operational Infrastructure Support Program.

Authors' Contributions

SMG, JDW, YJ, EEM, and SNT conceived the study; the original research protocol was jointly written by YJ, EEM, SMG, and JDW; all authors participated in the study design; YF was responsible for executing the study with assistance from BG, statistical analysis with input from JDW, EEM, and AF, and drafting the manuscript. All authors read, revised, and approved the final manuscript and the order of authorship reflects input.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study questionnaire.

[PDF File (Adobe PDF File), 421KB - jmir_v14i1e20_app1.pdf]

References

1. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006 Feb 1;163(3):197-203 [FREE Full text] [doi: 10.1093/aje/kwj036] [Medline: 16339049]
2. Blumberg SJ, Luke JV. Reevaluating the need for concern regarding noncoverage bias in landline surveys. Am J Public Health 2009 Oct;99(10):1806-1810. [doi: 10.2105/AJPH.2008.152835] [Medline: 19696381]
3. Garrett SK, Thomas AP, Cicuttini F, Silagy C, Taylor HR, McNeil JJ. Community-based recruitment strategies for a longitudinal interventional study: the VECAT experience. J Clin Epidemiol 2000 May;53(5):541-548. [Medline: 10812328]
4. Robinson JL, Fuerch JH, Winiewicz DD, Salvy SJ, Roemmich JN, Epstein LH. Cost effectiveness of recruitment methods in an obesity prevention trial for young children. Prev Med 2007 Jun;44(6):499-503. [doi: 10.1016/j.ypmed.2007.03.004] [Medline: 17475318]
5. Faden VB, Day NL, Windle M, Windle R, Grube JW, Molina BS, et al. Collecting longitudinal data through childhood, adolescence, and young adulthood: methodological challenges. Alcohol Clin Exp Res 2004 Feb;28(2):330-340. [Medline: 15112941]
6. Jones S, Fox S. Pew Internet project data memo: generations online in 2009. Washington, DC: Pew Internet & American Life Project; 2009 Jan 28. URL: http://pewinternet.org/~/media/Files/Reports/2009/PIP_Generations_2009.pdf [accessed 2011-10-28] [WebCite Cache ID 62lOhzr8T]
7. Lenhart A, Madden M, Hitlin P. Teens and technology: youth are leading the transition to a fully wired and mobile nation. Washington, DC: Pew Internet & American Life Project; 2005 Jul 27. URL: http://pewinternet.org/~/media//Files/Reports/2005/PIP_Teens_Tech_July2005web.pdf.pdf [accessed 2011-10-28] [WebCite Cache ID 62lOvEsgT]
8. Facebook.com. 2011. Statistics URL: http://www.facebook.com/press/info.php?statistics [WebCite Cache ID 62lPDvSk5]
9. Sevasti A. Memeburn. 2011 Jan 19. Facebook: Emerging Market Countries Surge Into the Top 10 URL: http://memeburn.com/2011/01/top-10-countries-on-facebook [accessed 2011-10-28] [WebCite Cache ID 62lPGLiMu]
10. Anonymous. The Australian Internet & Technology Report: Edition 12. Macquarie Park, NSW: The Nielsen Company; 2010.
11. Wallace B. McNair Ingenuity Research. 2011 Mar. Online privacy survey URL: http://hungrybeast.abc.net.au/sites/default/files/documents/1140M%20Online%20Privacy%20Survey%20-%20Report%20for%20Hungry%20Beast.pdf [accessed 2011-10-28] [WebCite Cache ID 62lPLAQLW]
12. 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]
13. 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] [Medline: 19073542]
14. Koo M, Skinner H. Challenges of internet recruitment: a case study with disappointing results. J Med Internet Res 2005;7(1):e6 [FREE Full text] [doi: 10.2196/jmir.7.1.e6] [Medline: 15829478]
15. Abdesslem FB, Parris I, Henderson T. Mobile experience sampling: reaching the parts of Facebook other methods cannot reach. In: Proceedings. 2010 Presented at: Privacy and Usability Methods Pow-Wow (PUMP); Sept 6, 2010; Dundee, UK URL: http://www.cl.cam.ac.uk/~fb375/publications/Ben-Abdesslem.PUMP10.pdf
16. Ahmed B, Jacob P, Allen F, Benowitz N. Attitudes and practices of hookah smokers in the San Francisco Bay Area. J Psychoactive Drugs 2011;43(2):146-152. [Medline: 21858960]
17. Battistella E, Kalyan S, Prior JC. Evaluation of methods and costs associated with recruiting healthy women volunteers to a study of ovulation. J Womens Health (Larchmt) 2010 Aug;19(8):1519-1524. [doi: 10.1089/jwh.2009.1751] [Medline: 20626268]
18. Jones SC, Magee CA. Exposure to alcohol advertising and alcohol consumption among Australian adolescents. Alcohol Alcohol 2011;46(5):630-637. [doi: 10.1093/alcalc/agr080] [Medline: 21733835]
19. Lord S, Brevard J, Budman S. Connecting to young adults: an online social network survey of beliefs and attitudes associated with prescription opioid misuse among college students. Subst Use Misuse 2011;46(1):66-76. [doi: 10.3109/10826084.2011.521371] [Medline: 21190407]
20. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775. [doi: 10.1093/ntr/ntq086] [Medline: 20530194]
21. Ryan T, Xenos S. Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. Comput Human Behav 2011;27(5):1658-1664. [doi: 10.1016/j.chb.2011.02.004]
22. Williams MT, Proetto D, Casiano D, Franklin ME. Recruitment of a hidden population: African Americans with obsessive-compulsive disorder. Contemp Clin Trials 2012 Jan;33(1):67-75. [doi: 10.1016/j.cct.2011.09.001] [Medline: 21983626]

23. Brickman-Bhutta C. Association of Religion Data Archives. 2009 Oct 27. Not by the Book: Facebook as Sampling Frame URL: http://www.thearda.com/workingpapers/download/Not%20by%20the%20Book%20-%20Bhutta.doc [accessed 2011-10-28] [WebCite Cache ID 62lQDNTaU]
24. Bull SS, Breslin LT, Wright EE, Black SR, Levine D, Santelli JS. Case study: An ethics case study of HIV prevention research on Facebook: the Just/Us study. J Pediatr Psychol 2011;36(10):1082-1092. [doi: 10.1093/jpepsy/jsq126] [Medline: 21292724]
25. Hoy MG, Milne G. Gender differences in privacy-related measures for young adult Facebook users. J Interact Advert 2010;10(2):28-45 [FREE Full text]
26. Facebook.com. 2011. Targeting Options URL: http://www.facebook.com/help/?page=203882222982239 [accessed 2011-11-28] [WebCite Cache ID 63Y1GnxOa]
27. Facebook.com. 2011. Campaign Cost and Budgeting URL: http://www.facebook.com/help/?page=219791638048948 [accessed 2011-11-28] [WebCite Cache ID 63Y1Rvj4Z]
28. Anonymous. Consent for Treatment and Confidentiality in Young People: Protecting Patients, Guiding Doctors. Melbourne: Medical Practitioners Board of Victoria; 2004.
29. Anonymous. 2006 Census Tables: 2006 Census of Population and Housing (Cat no 2068.0).: Australian Bureau of Statistics; 2006. URL: http://www.abs.gov.au/TableBuilder [accessed 2012-01-24] [WebCite Cache ID 64wIXYNx4]
30. Department of Health, Victoria, Australia. 2011 Jul 13. Victorian Population Health Survey 2008 URL: http://www.health.vic.gov.au/healthstatus/survey/vphs2008.htm [accessed 2011-10-28] [WebCite Cache ID 62lQcd0Q9]
31. Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia (Cat no 2033.0.55.001).: Australian Bureau of Statistics; 2008 Mar 26. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mediareleasesbytitle/87E66027D6856FD6CA257417001A550A?OpenDocument [accessed 2011-10-28] [WebCite Cache ID 62lQhfStF]
32. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. Am J Epidemiol 2005 Aug 1;162(3):199-200 [FREE Full text] [doi: 10.1093/aje/kwi188] [Medline: 15987728]
33. National Health and Medical Research Council, Australian Research Council, Australian Vice-Chancellors' Committee. National Statement on Ethical Conduct in Human Research. Canberra: Australian Government; 2007.
34. Anonymous. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser 2000;894:i-xii, 1-253. [Medline: 11234459]
35. Brown WJ, Dobson AJ, Bryson L, Byles JE. Women's Health Australia: on the progress of the main cohort studies. J Womens Health Gend Based Med 1999 Jun;8(5):681-688. [Medline: 10839654]
36. Vandelanotte C, Sugiyama T, Gardiner P, Owen N. Associations of leisure-time internet and computer use with overweight and obesity, physical activity and sedentary behaviors: cross-sectional study. J Med Internet Res 2009;11(3):e28 [FREE Full text] [doi: 10.2196/jmir.1084] [Medline: 19666455]
37. Schneider M, Dunton GF, Cooper DM. Media use and obesity in adolescent females. Obesity (Silver Spring) 2007 Sep;15(9):2328-2335 [FREE Full text] [doi: 10.1038/oby.2007.276] [Medline: 17890502]
38. Boyle RG, Enstad C, Asche SE, Thoele MJ, Sherwood NE. Evaluating strategies and costs to recruit smokeless tobacco users. Addict Behav 2007 Dec;32(12):3088-3092. [doi: 10.1016/j.addbeh.2007.06.005] [Medline: 17602843]
39. Sadler GR, Ko CM, Malcarne VL, Banthia R, Gutierrez I, Varni JW. Costs of recruiting couples to a clinical trial. Contemp Clin Trials 2007 Jul;28(4):423-432. [doi: 10.1016/j.cct.2006.11.006] [Medline: 17218166]
40. Rubin RR, Fujimoto WY, Marrero DG, Brenneman T, Charleston JB, Edelstein SL, DPP Research Group. The Diabetes Prevention Program: recruitment methods and results. Control Clin Trials 2002 Apr;23(2):157-171. [Medline: 11943442]
41. Communications Report 2009-10 Series: Report 2-Take-up and Use of Voice Services by Australian Consumers. Australian Communications and Media Authority. 2010 URL: http://www.acma.gov.au/webwr/_assets/main/lib310665/report_2_take_up_and_use_of_voice_comms.pdf [accessed 2012-01-26] [WebCite Cache ID 64zjJ0Il4]
42. Hu SS, Balluz L, Battaglia MP, Frankel MR. Improving public health surveillance using a dual-frame survey of landline and cell phone numbers. Am J Epidemiol 2011 Mar 15;173(6):703-711. [doi: 10.1093/aje/kwq442] [Medline: 21343246]
43. Wardropper S. Global Facebook Advertising Report: Q2 2011.: TBG Digital; 2011. URL: http://www.tbgdigital.com/TBG_Digital_Q2_2011_Global_Facebook_Marketing_Report.pdf [accessed 2011-11-28] [WebCite Cache ID 63Y4OA991]
44. Hoonakker P, Carayon P. Questionnaire survey nonresponse: a comparison of postal mail and Internet surveys. Int J Hum Comput Interact 2009;25(5):348-373. [doi: 10.1080/10447310902864951]
45. Madden M. Older adults and social media. Washington, DC: Pew Internet & American Life Project; 2010 Aug 27. URL: http://pewresearch.org/pubs/1711/older-adults-social-networking-facebook-twitter [accessed 2011-10-28] [WebCite Cache ID 62lRNU4Ih]
46. Hargittai E. Crooked Timber. 2009 Jul 8. Popularity of Facebook and MySpace changes, but SES differences in use persist URL: http://crookedtimber.org/2009/07/08/popularity-of-facebook-and-myspace-changes-but-ses-differences-in-use-persist/ [accessed 2011-10-28] [WebCite Cache ID 62lRRiDBN]

Abbreviations

ALSWH: Australian Longitudinal Study on Women's Health

BMI: body mass index
CI: confidence interval
PR: prevalence ratio
SEIFA: Socio-Economic Indexes For Areas

Edited by G Eysenbach; submitted 31.10.11; peer-reviewed by C Gaydos, M Koo, K Durant; comments to author 22.11.11; revised version received 14.12.11; accepted 06.01.12; published 01.02.12.

Please cite as:
Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, Gunasekaran B, Wark JD
Web-Based Recruiting for Health Research Using a Social Networking Site: An Exploratory Study
J Med Internet Res 2012;14(1):e20
URL: http://www.jmir.org/2012/1/e20/
doi:10.2196/jmir.1978
PMID:22297093

Original Paper

Broad Reach and Targeted Recruitment Using Facebook for an Online Survey of Young Adult Substance Use

Danielle E Ramo[1], PhD; Judith J Prochaska[1], PhD, MPH

Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Danielle E Ramo, PhD
Department of Psychiatry
University of California, San Francisco
401 Parnassus Avenue
Box TRC-0984
San Francisco, CA, 94143
United States
Phone: 1 415 476 7071
Fax: 1 415 476 7053
Email: danielle.ramo@ucsf.edu

Abstract

Background: Studies of tobacco use and other health behaviors have reported great challenges in recruiting young adults. Social media is widely used by young adults in the United States and represents a potentially fast, affordable method of recruiting study participants for survey research.

Objective: The present study examined Facebook as a mechanism to reach and survey young adults about tobacco and other substance use.

Methods: Participants were cigarette users, age 18-25 years old, living throughout the United States and recruited through Facebook to complete a survey about tobacco and other substance use. Paid advertising using Facebook's Ad program over 13 months from 2010 Feb 28 to 2011 Apr 4 targeted by age (18-25), location (United States or California), language (English), and tobacco- and/or marijuana-related keywords. Facebook approved all ads.

Results: The campaign used 20 ads, which generated 28,683,151 impressions, yielding 14,808 clicks (0.7% of targeted Facebook members), at an overall cost of $6,628.24. The average cost per click on an ad was $0.45. The success of individual ads varied widely. There was a rise in both clicks and impressions as the campaign grew. However, the peak for clicks was 3 months before the peak for ad impressions. Of the 69,937,080 accounts for those age 18-25 in the United States, Facebook estimated that 2.8% (n = 1,980,240) were reached through tobacco and marijuana keywords. Our campaign yielded 5237 signed consents (35.4% of clicks), of which 3093 (59%) met criteria, and 1548 (50% of those who met criteria) completed the survey. The final cost per valid completed survey was $4.28. The majority of completed surveys came from whites (69%) and males (72%). The sample averaged 8.9 cigarettes per day (SD 7.5), 3.8 years of smoking (SD 2.9), with a median of 1 lifetime quit attempts; 48% did not intend to quit smoking in the next 6 months.

Conclusions: Despite wide variety in the success of individual ads and potential concerns about sample representativeness, Facebook was a useful, cost-effective recruitment source for young-adult smokers to complete a survey about the use of tobacco and other substances. The current findings support Facebook as a viable recruitment option for assessment of health behavior in young adults.

(J Med Internet Res 2012;14(1):e28) doi:10.2196/jmir.1878

KEYWORDS

social media, Facebook, participant recruitment, young adult, tobacco

Introduction

Studies of tobacco use and other health behaviors have reported great challenges in recruiting young adults [1,2]. The Internet, increasingly used as a method to target and survey individuals about health-risk behaviors, may be a useful tool for reaching young adults. Compared to face-to-face interviews, Internet-based surveys can reach more potential respondents; enable inclusion of low-incidence or "hidden" population groups; enable rapid, convenient input by respondents; and reduce bias in response to sensitive, potentially stigmatizing topics [3-7]. A recent telephone survey of young adults age 18-29 in the United States indicates that almost all (93%) use the Internet [8]. Further, over the past decade, young adults have remained the age group most likely to go online.

Studies of Internet-based tobacco-cessation treatment have demonstrated high enrollment among general-aged adult participants through advertisements on Google or other search engines [9]. An intervention for smokeless tobacco (Chewfree.com) used advertisements on Google.com and generated 9155 clicks and 511 intervention participants at a cost of $6.70 per participant; advertisements on other search engines generated 363 participants (mean age, 34.5 years) in 15 months [10]. In another online smoking-cessation intervention (Quitnet [11]) advertisements on Google resulted in 28,296 clicks, producing 5557 eligible participants. Of these, 1489 gave informed consent and 764 (mean age 35.1 years) completed a baseline assessment in 6 weeks. In our previous work using the Internet to recruit young-adult smokers in survey research, advertisements across the Web yielded the largest proportion of recruited participants and completed surveys overall; however, Craigslist and an online sampling company were more successful at targeting young-adult smokers who went on to complete the survey and ultimately were the more cost-effective methods compared to Internet advertisements [12].

Social media is widely used by young adults in the United States. Nearly three quarters (72%) of online 18-29-year-olds use social-networking websites, with 45% doing so on a typical day [8]. Facebook, the largest social-networking website and second most popular website in the United States after Google [13], has more than 500 million users worldwide, half of whom use the site daily [14]. Facebook represents a potentially fast and affordable method of recruiting study participants for survey research, especially young adults who use the site in large numbers and on a frequent basis. Given the increasing interest in using social media for recruiting participants to research, the current study reports on the success of an ad campaign on Facebook, the leader in this space. We believe this is the first study to examine Facebook as a mechanism to reach and survey young adults about health behavior.

Methods

Participants

The target population was young-adult cigarette users, age 18-25 years, living throughout the United States. Individuals had to be English literate and smoke at least one cigarette in the past 30 days to be eligible for participation.

Facebook Recruitment Campaign

To reach young adults who had smoked recently, we paid for advertising using Facebook's Advertising (Ad) program over 13 months from 2010 Feb 28 to 2011 Apr 4. Our campaign involved creating advertisements that appeared on the pages of our target audience meeting the criteria of age (18-25), location (United States or California), language (English), and tobacco- and/or marijuana-related keywords that appeared in their Facebook profiles through listed interests, activities, education and job titles, pages they like, or groups to which they belong (eg, "cigarette," "nicotine," "blunt," "420"). At the time of this campaign, this was the only way Facebook ad could be targeted (ie, there was no way to target keywords to other areas of the Facebook profile). Facebook had to approve all ads based on the company's guidelines [15]. Only one ad type was available to advertisers at the time the campaign was launched. Ads included a short (eg, 2-word) headline, a picture, and a link to the study's survey website per Facebook's advertising size and word-count specifications. Facebook rejected an ad that targeted both tobacco and marijuana users through pictures. Therefore that ad made no impressions. We incurred a charge every time a user clicked on one of our ads.

On a daily basis, we could specify a spending limit for each ad and for the entire campaign. We, like other advertisers, could then specify the maximum amount we would be willing to pay for an ad (a "bid"). Then auctions determined the likelihood a given ad would be shown on pages of the target audience. Selection criteria included bid (the amount an advertiser is willing to pay), quality of an ad (including feedback an ad has received from users), and past performance [16]. For a given ad, Facebook suggests a "bid range" based on how much other advertisers would be willing to pay to reach the same target audience. This range can change over time based on both the ad space (other ads in the pool of ads) or an ad's performance. Our bids fluctuated over time in line with the bid range for a given ad. Facebook reports statistics on bids, impressions, clicks, and dollars spent on all ads in a campaign. Impressions are defined as a single time an ad is shown to a user, regardless of whether the user clicks on the ad. Clicks are when a user clicks a link in an ad. We used the Facebook-provided statistics indicating the success of each advertisement we ran and changed or stopped ads that were unsuccessful (ie, they rarely appeared, they received too few clicks, or they were too expensive). As such, we explored various picture and text options and determined the most successful ads based on impressions, clicks, and costs. Facebook, which normally restricts the reporting of its data, gave us permission, conveyed to our university's legal counsel, to publish these statistics.

Study Procedures

The Institutional Review Board approved study procedures, described in detail previously. Participants provided informed consent [12,17]. The online survey included basic demographics and a series of measures of smoking and other substance-use behaviors and thoughts about use. A Smoking Questionnaire assessed participants' years of smoking, prior quit attempts (lifetime and past year), and longest period of abstinence in a prior quit attempt [18]. The Smoking Stages of Change

Questionnaire [19] assessed motivation to quit, categorizing smokers into 1 of 3 preaction stages of change: precontemplation: no intention to quit within the next 6 months; contemplation: intention to quit within the next 6 months but no 24-hour quit attempt in the past year; preparation: intention to quit within the next month and a 24-hour quit attempt in the past year. The Thoughts about Abstinence form [20] assessed desire to quit, anticipated success with quitting, and perceived difficulty with abstinence (each rated on a scale of 1 to 10). Participants were required to answer all questions before they could continue to the next page of the survey and could quit the survey at any time. Computer Internet Protocol (IP) addresses were tracked and multiple entries were not accepted from the same computer. Data were deemed invalid and excluded from analyses if (1) there was a discrepancy in data from duplicate questions (eg, date of birth; $n = 215$), (2) respondents reported the same contact email address across multiple survey entries ($n = 50$), or (3) the data were clearly invalid (eg, every entry was the same across the entire survey; $n = 28$). At the end of the recruitment period, each completed survey entry associated with a valid email address was entered into a drawing to win a $400 gift certificate to Apple stores or a $25 gift certificate to a national or online store.

Results

Advertising Campaign

During the 13-month campaign, our ads made 28,683,151 impressions, yielding 14,808 clicks, at an overall cost of $6,628.24. The average cost per click on an ad was $0.45. Twenty different ads were run. Of those, 14 were targeted throughout the United States and 6 specifically to California; 7 asked participants if they "ever smoke," 7 asked if they "ever smoked cigarettes," and 6 asked if they "smoked recently;" 8 had a cartoon picture, including lit cigarettes (n = 3) and a cigarette pack (n = 5); and 12 had a realistic picture, including a lit cigarette (n = 3), a cigarette pack (n = 3), multiple cigarettes (n = 3), and a person smoking a cigarette (n = 3).

The success of individual ads varied widely. Figure 1 illustrates, as an example, 2 sets of 3 ads deemed highly successful, moderately successful, or unsuccessful based on their campaign statistics. The most successful ad in our campaign had over 8 million impressions, which yielded over 5000 clicks, and cost $0.38 per click (Figure 1a). Other ads varied in the number of clicks and costs per click (eg, Figure 1b). Some ads made impressions but were not clicked on at all. For these we did not incur any charges (Figure 1c).

Figure 2 presents the rate of impressions and clicks compared to campaign costs for the entire campaign over 13 months. There was a rise in both clicks and impressions as the campaign grew. However, the peak for clicks was in the summer of 2010, which was 3 months before the peak for ad impressions (October 2010). There was a large drop-off in impressions between October 2010 and January 2011, without a corresponding shift in clicks or costs, highlighting that the clicks were coming from a few consistently successful ads throughout that period.

Recruitment Results

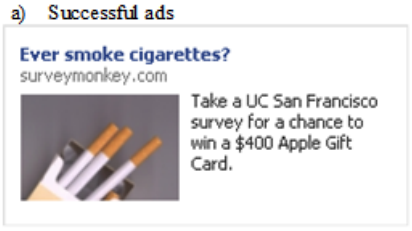
Figure 3 summarizes the numbers of potential Facebook accounts reached through various target characteristics, the clicks our ads received, and the sample who reached and completed our survey. Of the 69,937,080 Facebook accounts registered to individuals age 18-25 in the United States, 2.8% (n = 1,980,240) had profiles with tobacco- and marijuana-related keywords (our target population). Our campaign received 14,808 clicks (0.7% of potential accounts reached through the campaign), which yielded 5237 signed consents, of which 3093 (n = 59%) met study-inclusion criteria, and 1548 (50% of those who met criteria) completed the survey. The final cost per valid, completed survey was $4.28.

Participant Characteristics

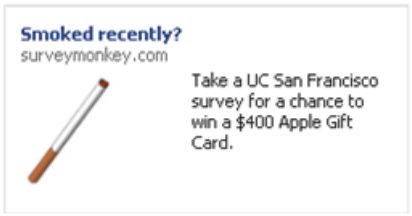
Table 1 presents sociodemographic and tobacco-use characteristics of the sample that completed the survey. The majority of participants were male and white. All four US census regions were represented, with the Northeast having the lowest representation (20.2%). The sample averaged 8.9 cigarettes per day (SD 7.5) and 3.8 years of smoking (SD 2.9). Sixty-three percent reported a lifetime quit attempt of at least 24-hours' duration. At the time of survey completion, 48% were not intending to quit smoking in the next 6 months (ie, precontemplation stage of change); 30% were contemplating quitting in the next 6 months, but not in the next 30 days (ie, contemplation); and 23% were preparing to quit in the next 30 days and reported a quit attempt in the past year (ie, preparation).

Figure 1. Examples of two successful (1a), moderately successful (1b), and unsuccessful (1c) advertisements from the Facebook campaign based on ad statistics. Reported Facebook-suggested bids are from the last day of the campaign.

a) Successful ads



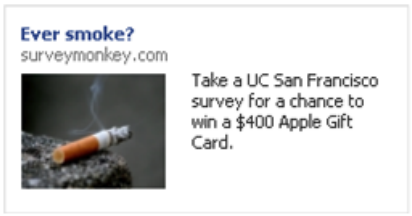
Facebook Suggested Bid: $0.80 – $1.27
Total Impressions: 8,492,286
Total Clicks: 5,002
Cost per click: **$0.38**
Cost per thousand impressions: $0.22
Total spent: $1,885.52
Response rate (clicks/impressions): 0.06%



Facebook Suggested Bid: $0.91 – $1.38
Total Impressions: 6,531,333
Total Clicks: 3,840
Cost per click: **$0.39**
Cost per thousand impressions: $0.23
Total spent: $1504.18
Response rate (clicks/impressions): 0.06%

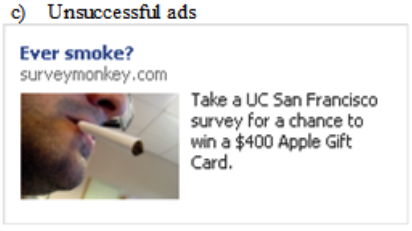
b) Moderately Successful ads

Facebook Suggested Bid: $1.28 – $2.03
Total Impressions: 1,418,704
Total Clicks: 524
Cost per click: **$0.61**
Cost per thousand impressions: $0.22
Total spent: $318.75
Response rate (clicks/impressions): 0.04%

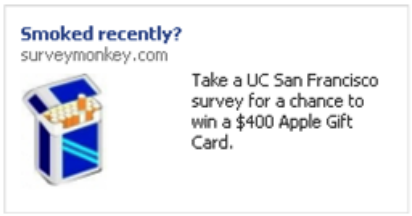


Facebook Suggested Bid: $1.36 – $2.07
Total Impressions: 2,451,283
Total Clicks: 963
Cost per click: **$0.56**
Cost per thousand impressions: $0.22
Total spent: $543.33
Response rate (clicks/impressions): 0.04%

c) Unsuccessful ads



Facebook Suggested Bid: $2.02 – $3.05
Total Impressions: 28,231
Total Clicks: 5
Cost per click: **$0.70**
Cost per thousand impressions: $0.12
Total spent: $3.51
Response rate (clicks/impressions): 0.02%



Facebook Suggested Bid: $1.66 – $2.52
Total Impressions: 5,250
Total Clicks: 0
Cost per click: **$0.00**
Cost per thousand impressions: $0.00
Total spent: $0.00
Response rate (clicks/impressions): 0.00%

Figure 2. Total monthly impressions and clicks (left axis) compared to average monthly dollars spent (right axis) across the 13-month Facebook recruitment campaign.

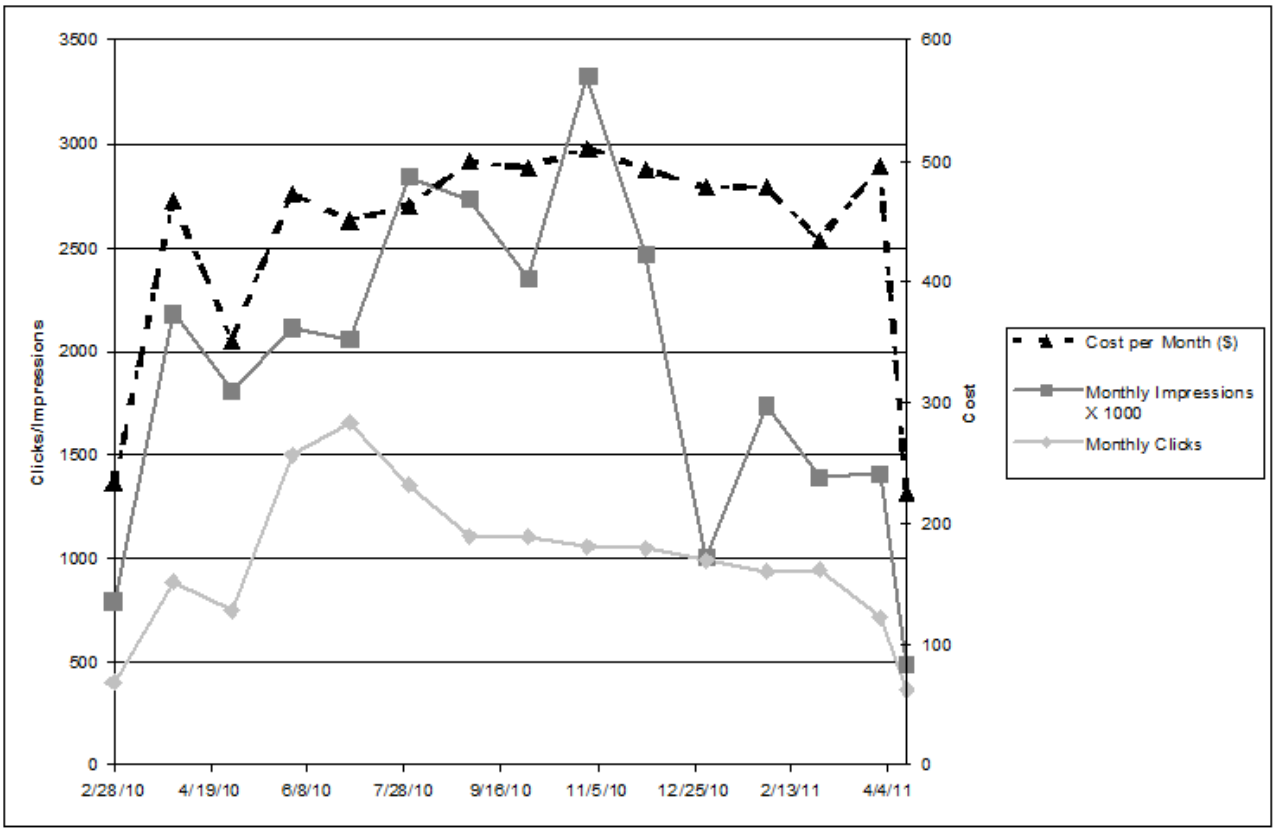


Figure 3. Facebook ad campaign reach and recruitment process.

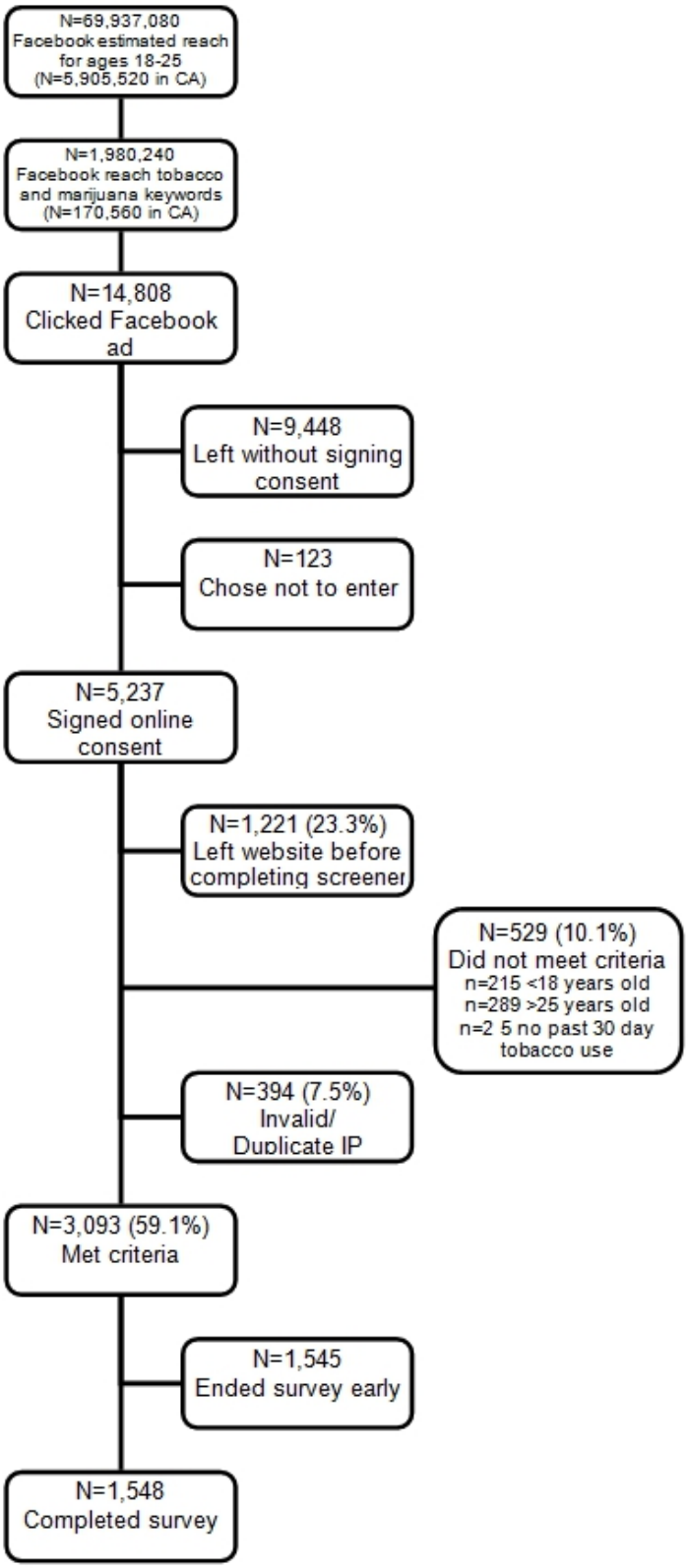


Table 1. Demographic and smoking characteristics of young adults who completed the survey (N = 1548)

General characteristics	% or mean (SD)
Gender, %	
Female	30.6
Male	68.9
Transgender	0.5
Age in years, mean (SD)	20.3 (2.0)
Race/ethnicity, %	
African-American/black	2.5
Asian/Pacific Islander	3.6
White	71.7
Hispanic/Latino	6.1
Other	16.1
Employment status, %	
Full-time	28.7
Part-time	17.0
Unemployed/homemaker	24.0
Student	30.3
Years of education, mean (SD)	13.0 (1.9)
Annual family income, %	
Less than $20,000	25.3
$21,000-$40,000	20.0
$41,000-$60,000	14.8
$61,000-$80,000	11.2
$81,000-$100,000	9.4
Over $100,000	19.4
Subjective social status, mean (SD)	5.8 (1.9)
Region, %	
Northeast	20.2
Midwest	26.9
South	27.1
West	25.8
Number of cigarettes smoked per day, mean (SD)	8.9 (7.5)
Years smoking, mean (SD)	3.8 (2.9)
Prior quit attempts in lifetime, mean (SD)	8.5 (26.6)
Quit attempts in past year: median (interquartile range)	1.0 (3)
Stage of change, %	
Precontemplation	47.5
Contemplation	29.6
Preparation	22.9
Desire to quit, mean (SD)	5.3 (3.0)
Expected success, mean (SD)	6.0 (2.9)
Expected difficulty, mean (SD)	6.5 (2.8)

Discussion

Overall, Facebook was a successful recruitment source for young-adult smokers to complete a survey about tobacco and other substance use. At $4.28 per completed survey, this method proved much more affordable than other methods we used in previous survey research with this population, especially given the minimal staff time to design and monitor the campaign. In our prior work, the cost per completed survey was $42.77 for an Internet marketing company to place ads throughout the web and $19.24 for a survey sampling company to email survey announcements to young adults [12]. Traditional recruitment mechanisms, such as newspaper or radio advertising or hired survey sampling companies, tend to be far more expensive and take more time for participant screening. Other Internet-based recruitment mechanisms that have proven successful at recruiting for smoking-cessation interventions, such as Google's AdWords program [9,10], are not as easy to use for targeting a specific demographic population. As a result, such mechanisms require users to screen participants more thoroughly and to validate data generated through these recruitment methods.

Facebook is a useful mechanism for recruiting young research participants in a field that has shown past challenges. Nearly half (48%) of those completing the survey were unmotivated to quit smoking within the next 6 months. Smokers unmotivated to quit may be particularly challenging to recruit into face-to-face research studies, further highlighting the value of online methods. The current findings indicate that Facebook, compared to other mechanisms, may be particularly effective at recruiting smokers unmotivated to quit in the near future as the proportion was higher than that in our prior research of online recruitment of young-adult smokers [12]. Young adults are conducting daily communications at an increasing rate via mobile methods, rather than by face-to-face encounters or by telephone at home, making traditional methods of recruitment and assessment increasingly obsolete [8]. Researchers desiring to understand health behavior of young adults (especially stigmatized behavior) should consider social media as a viable option for recruitment.

The campaign's success was predicated upon Facebook's approval of each advertisement. The advertising policy strictly forbids the sale of or reference to tobacco products as well as mention of illegal activity [15]. To obtain Facebook approval of our tobacco ads, we provided evidence that we represented an academic research study. However, an ad with a picture of a marijuana leaf on it was not approved. Facebook's wide reach makes it a useful way to target hard-to-reach populations; however, to adhere to Facebook's guidelines, studies dealing with illegal activity, such as underage alcohol use or illicit drug use, may be required to use more neutral words and pictures, and thus may reach a wider population than their intended subject pool.

A limitation to recruitment through Facebook is that the representativeness of the sample cannot be fully determined. At the time of recruitment, Facebook reported the number of total Facebook accounts potentially targeted by any given set of demographic characteristics or keywords; however, there was no ability to compare this larger population to the eventual sample generated through an ad campaign. At the time the campaign was implemented, there was also no way to determine how many individuals within the population were exposed to at least one advertisement. It is possible that Facebook will choose to make this information available but it is at their discretion. The data that Facebook did provide for each ad (eg, impressions, clicks, costs) made it simple to continuously evaluate and optimize our campaign.

The research presented here used a cross-sectional survey design. There is limited knowledge about Facebook's success at reaching and retaining young adults in longitudinal treatment-outcome studies or clinical trials. Given the high proportion of recruited smokers unmotivated to quit, tobacco-treatment models that do not require immediate cessation (eg, stage-tailored interventions, smoking reduction) may be particularly appropriate, although this should be confirmed. As a common method of communication for young people all over the world, social media represents a useful strategy that can be leveraged for research to find and engage potentially hard-to-reach populations.

Acknowledgments

This study was supported by an individual postdoctoral fellowship award from the California Tobacco-Related Diseases Research Program (18FT-0055; Danielle E. Ramo, principal investigator), and an institutional training grant (T32 DA007250; James L. Sorensen, principal investigator) and center grant from the National Institute on Drug Abuse (P50 DA09253; Joseph Guydish, principal investigator). The preparation of this manuscript was supported in part by a career development award from National Institute on Drug Abuse (K23 DA018691; Judith J. Prochaska, principal investigator), a research project grant from the National Institute of Mental Health (R01 MH083684, Judith J. Prochaska, principal investigator), and a research award from the California Tobacco-Related Disease Research Program (#13-KT-0152). None of the funding bodies had any role in reviewing or approving the manuscript for publication. The authors thank Howard Liu for his work on the study and his help preparing figures for this manuscript.

Conflicts of Interest

None declared.

References

1. Bost ML. A descriptive study of barriers to enrollment in a collegiate health assessment program. J Community Health Nurs 2005;22(1):15-22. [doi: 10.1207/s15327655jchn2201_2] [Medline: 15695193]
2. Davies J, McCrae BP, Frank J, Dochnahl A, Pickering T, Harrison B, et al. Identifying male college students' perceived health needs, barriers to seeking help, and recommendations to help men adopt healthier lifestyles. J Am Coll Health 2000 May;48(6):259-267. [doi: 10.1080/07448480009596267] [Medline: 10863869]
3. Schonlau M, Zapert K, Simon LP, Sansad KH, Marcus SM, Adams J, Spranca M, Kan H, Turner R, Berry SH. A comparison between responses from a propensity-weighted web survey and an identical RDD survey. Soc Sci Comput Rev 2004;22(1):128-138. [doi: 10.1177/0894439303256551]
4. Cantrell MA, Lupinacci P. Methodological issues in online data collection. J Adv Nurs 2007 Dec;60(5):544-549. [doi: 10.1111/j.1365-2648.2007.04448.x] [Medline: 17973718]
5. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. J Epidemiol Community Health 2003 Jan;57(1):68-73 [FREE Full text] [Medline: 12490652]
6. Schonlau M, van Soest A, Kapteyn A. Are "Webographic" or attitudinal questions useful for adjusting estimates from Web surveys using propensity scoring? Surv Res Methods 2007;1(3):155-163.
7. McCabe SE, Boyd CJ, Couper MP, Crawford S, D'Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. J Stud Alcohol 2002 Nov;63(6):755-761. [Medline: 12529076]
8. Lenhart A, Purcell K, Smith A, Zickuhr K. Social Media and Young Adults. Washington, DC: Pew Internet & American Life Project; 2010 Feb 03 [FREE Full text]
9. 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] [Medline: 19640833]
10. 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]
11. 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]
12. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the Internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775. [doi: 10.1093/ntr/ntq086] [Medline: 20530194]
13. Alexa. 2010. Top sites in United States URL: http://www.alexa.com/topsites/countries/US [accessed 2011-06-23] [WebCite Cache ID 5zfMvi1aP]
14. Facebook. 2010. Statistics URL: http://www.facebook.com/press/info.php?statistics [accessed 2010-11-06] [WebCite Cache ID 5zlUcRcbL]
15. Facebook. 2011. Facebook advertising guidelines URL: http://www.facebook.com/ads/mistakes.php#/ad_guidelines.php [accessed 2011-04-28] [WebCite Cache ID 5zlUfrgyW]
16. Facebook. 2011. Delivery: ensuring your ads receive impression and clicks URL: http://www.facebook.com/adsmarketing/index.php?sk=delivery [accessed 2011-04-27] [WebCite Cache ID 5zlVZ6vCg]
17. Ramo DE, Hall SM, Prochaska JJ. Reliability and validity of self-reported smoking in an anonymous online survey with young adults. Health Psychol 2011 Nov;30(6):693-701. [doi: 10.1037/a0023443] [Medline: 21574709]
18. Hall SM, Tsoh JY, Prochaska JJ, Eisendrath S, Rossi JS, Redding CA, et al. Treatment for cigarette smoking among depressed mental health outpatients: a randomized clinical trial. Am J Public Health 2006 Oct;96(10):1808-1814. [doi: 10.2105/AJPH.2005.080382] [Medline: 17008577]
19. 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]
20. Hall SM, Havassy BE, Wasserman DA. Commitment to abstinence and acute stress in relapse to alcohol, opiates, and nicotine. J Consult Clin Psychol 1990 Apr;58(2):175-181. [Medline: 2335634]

Edited by G Eysenbach; submitted 29.06.11; peer-reviewed by M Blankers, J Grohol; comments to author 13.09.11; revised version received 05.12.11; accepted 02.01.12; published 23.02.12.

Please cite as:
Ramo DE, Prochaska JJ
Broad Reach and Targeted Recruitment Using Facebook for an Online Survey of Young Adult Substance Use
J Med Internet Res 2012;14(1):e28
URL: http://www.jmir.org/2012/1/e28/
doi:10.2196/jmir.1878
PMID:22360969

Original Paper

Matrix Analysis of the Digital Divide in eHealth Services Using Awareness, Want, and Adoption Gap

Te-Hsin Liang[1], PhD

Department of Statistics and Information Sciences, College of Management, Fu Jen Catholic University, New Taipei County, Taiwan

Corresponding Author:
Te-Hsin Liang, PhD
Department of Statistics and Information Sciences
College of Management
Fu Jen Catholic University
Room 240, No 510, Zhongzheng Rd
Xinzhuang District
New Taipei County, 242
Taiwan
Phone: 886 229052974
Fax: 886 229052191
Email: stat1013@mail.fju.edu.tw

Abstract

Background: The digital divide usually refers to access or usage, but some studies have identified two other divides: awareness and demand (want). Given that the hierarchical stages of the innovation adoption process of a customer are interrelated, it is necessary and meaningful to analyze the digital divide in eHealth services through three main stages, namely, awareness, want, and adoption.

Objective: By following the three main integrated stages of the innovation diffusion theory, from the customer segment viewpoint, this study aimed to propose a new matrix analysis of the digital divide using the awareness, want, and adoption gap ratio (AWAG). I compared the digital divide among different groups. Furthermore, I conducted an empirical study on eHealth services to present the practicability of the proposed methodology.

Methods: Through a review and discussion of the literature, I proposed hypotheses and a new matrix analysis. To test the proposed method, 3074 Taiwanese respondents, aged 15 years and older, were surveyed by telephone. I used the stratified simple random sampling method, with sample size allocation proportioned by the population distribution of 23 cities and counties (strata).

Results: This study proposed the AWAG segment matrix to analyze the digital divide in eHealth services. First, awareness and want rates were divided into two levels at the middle point of 50%, and then the 2-dimensional cross of the awareness and want segment matrix was divided into four categories: opened group, desire-deficiency group, perception-deficiency group, and closed group. Second, according to the degrees of awareness and want, each category was further divided into four subcategories. I also defined four possible strategies, namely, hold, improve, evaluate, and leave, for different regions in the proposed matrix. An empirical test on two recently promoted eHealth services, the digital medical service (DMS) and the digital home care service (DHCS), was conducted. Results showed that for both eHealth services, the digital divides of awareness, want, and adoption existed across demographic variables, as well as between computer owners and nonowners, and between Internet users and nonusers. With respect to the analysis of the AWAG segment matrix for DMS, most of the segments, except for people with marriage status of Other or without computers, were positioned in the opened group. With respect to DHCS, segments were separately positioned in the opened, perception-deficiency, and closed groups.

Conclusions: Adoption does not closely follow people's awareness or want, and a huge digital divide in adoption exists in DHS and DHCS. Thus, a strategy to promote adoption should be used for most demographic segments.

(J Med Internet Res 2012;14(1):e11) doi:10.2196/jmir.1670

KEYWORDS

Consumer behavior process; digital divide; eHealth; innovation adoption process

Introduction

Health care organizations are beginning to use the Internet in reaching a large part of the population in a cost-effective manner [1]. Several hundred thousand websites worldwide with varying qualities of health information are accessed and used by consumers and professionals [2]. The diffusion of broadband, wireless, and mobile Internet [3] has likewise influenced the traditional behavior of consumer activities, even in health care, thereby bringing about various social benefits. eHealth has changed the way health care is delivered and practiced [4]. For patients, who can also be viewed as consumers, eHealth presents an opportunity to change their relationship with providers, such as doctors and nurses [5]. The adoption of eHealth innovations can have a significant impact on the wellness of communities and populations [6].

eHealth services have improved access to health care in rural [7,8], suburban [9,10], and urban areas [11]. eHealth is particularly useful in linking specialists in academic health centers with health care professionals in areas short of facilities for patient care [12]. Following the rapid development of broadband Internet access services, the digital divide across demographic variables has become a huge social issue [13]. Affordable, high-speed wireless Internet access can be provided in rural and remote areas, bridging the gap between health care service and customers [14]. However, the availability of Internet access might cause another digital divide in eHealth between Internet users and nonusers, as well as between computer owners and nonowners. In fact, the digital divide in access to Internet technology has already caused inequalities in terms of health care [15].

In the last 10 years, researchers have begun discussing customer acceptance of eHealth services using the technology acceptance model [16,17] and the theory of planned behavior [16]. However, previous studies have simply discussed eHealth service adoption from the system design and improvement side, and scarcely explored the adoption of specific eHealth services. Studies examining the digital divide in eHealth services from a hierarchy-type viewpoint, such as which customers are adopting a new product or service, are rare. Therefore, the two main aims of this study are as follows: (1) from the customer segment viewpoint, to propose a new matrix analysis of the digital divide using the awareness, want, and adoption gap ratio (AWAG segment matrix), and thereafter compare the digital divide among different groups, and (2) to conduct an empirical study on specific eHealth services and show the practicability of the proposed matrix analysis.

Methods

Literature Review and Proposed Hypotheses

The digital divide relates not only to Internet access but also to the existence of a gap between people who can effectively use new information and communication tools, such as the Internet, and those who cannot [18]. The digital divide usually refers to access or usage; however, some studies have also identified two other divides: awareness [19] and demand (want) [20,21]. Barriers to the emergence of an equitable information society have led to the existence of the digital divide [22].

More differentiated use of the Internet across varying segments of a given population may result in the digital divide [22,23]. Moreover, demographic variables and socioeconomic status are factors influencing the digital divide [24,25]. Previous studies have indicated that the digital divide across demographic variables, including gender [20,26,27], age [20,26,28,29], education [26-31], income [26,27,29,32], marital status [26,30], geographic area [13,20,29,31,33], and ethnicity [20,31,34], are significant. Low-income [35] and elderly people, and those living in rural areas constitute the digitally underserved population [20,33], whereas people with higher education levels or of younger age are considered the digitally leading population [28,30]. Most studies have indicated that gender is no longer an influential factor in the digital divide [26,29,32,36]. However, some studies have asserted that, whereas males are most likely to access the Internet and play online games [26,37,38], females are most likely to use eHealth services [20]. Divorced people are more isolated than those who are married; this may contribute to a tendency not to use eHealth services [30].

The availability of a home computer is another factor used to predict an individual's ability to access the Internet [24]. The ability to use a computer has been found to be associated with access to health-related information from the Internet [30]. People who are ill and have computer and Internet access desire specific information and may be more receptive to health information on managing their diseases [30,36,39].

Previous studies have shown that certain demographic variables and computer and Internet access are factors causing the digital divide, and that such a divide usually entails access or usage. Some studies have also identified two other divides: awareness [19] and demand (want) [20,21]. Thus, this research proposed three main hypotheses, with each having three subhypotheses, as follows:

H1: There exists an awareness divide in eHealth services across certain demographic variables, computer ownership, and Internet access.

H2: There exists a want divide in eHealth services across certain demographic variables, computer ownership, and Internet access.

H3: There exists an adoption divide in eHealth services across certain demographic variables, computer ownership, and Internet access.

The earliest and most well-known consumer purchasing decision process is attention–interest–desire–action, first proposed in the late 1800s and early 1900s [40,41]. Attention–interest–desire–action states that salespeople have to attract attention (cognition), maintain interest, and create desire (affect), leading to action (conation) [6,42]. Different models of consumer purchasing decisions consist of a sequence of mental stages or levels that consumers experience throughout the decision process [43-48]. Different studies have their own viewpoints, but most hierarchical models include six hierarchical stages: awareness, knowledge, liking, preference, conviction, and purchase. Some studies [47,49] have summarized the hierarchical stages of the consumer purchasing decision model

into three stages: awareness, interest, and final decision. In the first stage, awareness, the consumer knows that an alternative exists but may not have the interest or sufficient information to understand its possible benefits. In the second stage, interest, the consumer is aware, develops some interest, and hence decides to learn more about the product. In this stage, the wants of consumers are singled out. In the last stage, final decision, the consumer takes an observable action, which is the purchase of a good or service or the sustained adoption of an innovation.

Some studies have mentioned that probabilities can be associated with the stages of the hierarchical models to show the ultimate behavioral impact of promotion [50-52]. Therefore, when evaluating the digital divide in eHealth services, the percentages or probabilities of awareness, want, and adoption, corresponding to the three main stages of purchasing decision, should be measured. Consumers' want for eHealth services should also be created typically through promotion and education. However, the awareness of an eHealth service does not necessarily translate into choice or usage if there is a shortage of want. In other words, adoption does not occur if there is a shortage of awareness and want. The adoption rate of an e-service should be highly related to the corresponding awareness and want rates of individuals [53]. Thus, the following hypotheses were proposed:

H4: The adoption rate of a given eHealth service is bound to consumers' corresponding awareness rate.

H5: The adoption rate of a given eHealth service is bound to consumers' corresponding want rate.

From the end user's viewpoint, Dixon [54] proposed the information technology adoption model (ITAM), which was compiled from several technology adoption models and incorporates end-user satisfaction. ITAM is based on a triangular structure of design–implementation–evaluation: it demonstrates the chicken-and-egg connection between the process of innovation design, and its implementation and evaluation. Referring to the concept of ITAM, the movement of a product or information between two subjects distinguishes technology push from consumer pull. Technology push, which is similar to the chicken analogy of ITAM, is mainly driven by research and development activities; and consumer pull, which is similar to the egg analogy of ITAM, is driven by external market forces. In the market, officers or suppliers *push* new products toward consumers. Meanwhile, consumers *pull* the goods or information they demand. A push marketing strategy is used when there is development or improvement on a product unknown to consumers. Given that there is no consumer demand in a product launch, the product and the information are pushed to consumers by distribution and promotion [55,56].

In a pull health care system, the patient requests the product and pulls it through the delivery channel [57]. Taking the online registration service of outpatients as an example, in the beginning, most patients did not request the service. The service was simply pushed to them through promotion by hospitals and the government. The patients were then made aware of such a service, and they considered whether they liked or needed it. Following an increased awareness of the online registration service for outpatients, designers have developed new functions needed by patients. This suggests that the demand from patients pulled the supply, as well as the corresponding improvements brought about by heightened awareness. Therefore, awareness and want gradually rise through the cycle of technology push and consumer pull. In general, want is initiated and raised when awareness spreads. The adoption rate is raised when the awareness of and want for a given e-service spread. In other words, the want rate should be lower than the awareness rate. However, according to the above discussion of pull and push, for some consumer segments, the want rate for new and innovative e-services is not necessarily always lower than their corresponding awareness rate. Thus, the following hypotheses were proposed:

H6: The want rate for a given eHealth service is not necessarily bound to consumers' corresponding awareness rate.

H6-1: The want rate for a given eHealth service is bound to consumers' corresponding awareness rate.

H6-2: The want rate for new and innovative e-services may be greater than consumers' corresponding awareness rate.

Generally speaking, there is no adoption if there is no awareness. Higher awareness may bring higher want rates, but the intention of using some eHealth services will be low if there is a shortage of want. However, people having the potential need for an eHealth service will easily pay attention to the promotion and receive the information and, as such, may have a higher awareness rate than those who are not in need of the service. Thus, the following three hypotheses were proposed:

H7: Want rate, given awareness for each consumer segment, is higher than want rate with unawareness.

H8: Adoption rate, given want for each consumer segment, is higher than adoption rate without want.

H9: Awareness rate, given want for each consumer segment, is higher than awareness rate without want.

AWAG Segment Matrix

According to H4 to H9, awareness and want have interactive influences on adoption rate. Therefore, when evaluating the digital divide in some e-services, the corresponding awareness and want rates should be considered separately. Based on the technology adoption lifecycle (bell curve), with a combination of innovators and early majority stages, the four types of adopters are segmented by three slightly adjusting adoption life cycle cumulative rates of 15%, 50%, and 85% [42]. The awareness and want rates should be high for innovators or early adopters—that is, following the rise of the innovation level, the adoption life cycle cumulative rates should move from low to high. The present study used the above three cumulative rates to segment and position groups.

First, awareness and want rates were divided into two levels at the middle point of 50%, and then the 2-dimensional cross of the AWAG segment matrix was divided into four categories: opened group, desire-deficiency group, perception-deficiency group, and closed group. Second, using the cumulative rate of 15% or 85%, each category was further divided into four subcategories. In the awareness–want segment matrix (Figure

1), the location of a group indicates its awareness and want rates for an eHealth service.

People categorized under the opened group are open to innovation. They are keen on seeking new information and are always on the lookout for something new and innovative. On the other hand, those who are categorized under the closed group are closed minded when it comes to innovation, and they lag behind in receiving new information. They are not interested in innovation and, as such, they usually resist trying something new. People categorized under the desire-deficiency group lack desire for innovation. Although they receive new information early, they are usually not interested in innovation and may resist trying something new. Meanwhile, those under the perception-deficiency group lack perception for innovation. Although they lag behind in receiving new information, they are still interested in innovation and always intend to try something new and innovative.

Each of the above four groups was further divided into four subgroups, according to the degrees of awareness and want, and based on a cumulative rate of 15% or 85%. For the opened group and closed group, the subgroups were strong, awareness-bias, want-bias, and generic. For the desire-deficiency group and perception-deficiency group, the subgroups were strong, generic, and want-bias or awareness-bias. The strong subgroup is the most open, closed, perception-deficient, or desire-deficient group. The generic subgroup is the least open, closed, perception-deficient, or desire-deficient group. There is some room to raise awareness for the awareness-bias subgroup and some room to raise want for the want-bias subgroup.

In the AWAG segment matrix, awareness and want are on the same level for four groups: strong opened group, generic opened group, generic closed group, and strong closed group. In these four groups, the awareness rate corresponds to the want rate. However, the opened degree for innovation decreases from left-up to right-down. For example, people under the strong opened group are innovators with the most open minds. Most of them already know about some innovations or new services, and they are full of want. On the opposite side, people under the strong closed group are laggards with the most closed minds. Most of them do not know or care about innovation or new services, and they are lacking in want.

Groups located on the left-down side of the downward-sloping 45° line have awareness rates greater than want rates. The groups located on the opposite side have inversed characteristics. The larger the distance beyond the 45° line, the greater the bias between awareness and want. For example, the group in the farthest left-down area is the strong desire-deficiency group. People belonging to the strong desire-deficiency group may not be the target of innovation. Although they have high awareness, they are short of want. Thus, any innovation promotion will not drive them to do something, and any promotion budget allocated to this group may be wasted. People belonging to the strong perception-deficiency group located at the farthest right-up, although high in want of innovation, are seriously ignored or may not have the capability to get information. Thus, they do not receive enough information on innovation. This group should be prioritized first, and more promotion efforts should be exerted on them.

The present study defined four possible strategies: hold, improve, evaluate, and leave [53,58,59]. The hold strategy maintains the good work for innovators, early adopters, and the early majority. The improve strategy includes three types of strategies: spread, create, and raise. The spread strategy promotes awareness by adjusting the communication channel or method for a segment. The create strategy identifies and forms new wants for a specific group. The raise strategy raises awareness or want for a segment. The evaluate strategy re-evaluates wants for a segment, and then further chooses from the leave or improve strategy. The leave strategy suggests not taking action in some specific groups because they are nontarget markets and should be left alone. Each group and the corresponding strategies and actions suggested are shown in Table 1.

Table 1. Prescriptions of the awareness, want, and adoption gap ratio (AWAG) segment matrix

Category	Subcategory	Strategy	Action	Current target market
Opened group	Strong	Hold	Keep up the good work.	Primary
	Awareness-bias	Hold and improve (raise)	Keep up the good work and keep raising the awareness.	Secondary
	Want-bias	Hold and improve (raise)	Keep up the good work and keep raising the want.	Tertiary
	Generic	Hold and improve (raise)	Keep up the good work and keep raising the awareness and want.	Tertiary
Closed group	Strong	Evaluate then leave or evaluate then improve (spread and create)	Evaluate the potential of the segment then choose an action between "maintain status quo" and "keep spreading the awareness or creating the want."	Nontarget
	Awareness-bias	Evaluate then leave or evaluate then improve (spread)	Evaluate the potential of the segment then choose an action between "maintain status quo" and "keep spreading the awareness."	Nontarget
	Want-bias	Evaluate then leave or evaluate then improve (create)	Evaluate the potential of the segment then choose an action between "maintain status quo" and "keep creating the want."	Nontarget
	Generic	Evaluate then leave or evaluate then improve (raise)	Evaluate the potential of the segment then choose an action between "maintain status quo" and "keep raising the awareness or want."	Nontarget
Desire-deficient group	Strong	Evaluate then leave or evaluate then improve (create)	Evaluate the potential of the segment then choose an action between "maintain status quo" and "keep creating the want."	Nontarget
	Want-bias	Evaluate then leave or evaluate then improve (create)	Evaluate the potential of the segment then choose an action between "maintain status quo" and "keep creating the want."	Nontarget
	Generic	Improve (create)	Keep creating the want.	Nontarget
Perception-deficient group	Strong	Improve (spread)	Keep spreading the awareness.	Potential target
	Awareness-bias	Improve (spread)	Keep spreading the awareness.	Potential target
	Generic	Improve (spread)	Keep spreading the awareness.	Potential target

In using the AWAG segment matrix, managers should first re-evaluate the awareness and then the wants of some segments. If a segment is found to have low awareness, some promotion activities should be carried out, and follow-up should be conducted to raise the want. The e-services at the bottom right area should be pulled to the top left area, step by step, if possible or necessary. The suggested improvement direction for each group is shown in Figure 2. For example, the government of Taiwan has promoted the long-term management of physiological conditions since 2006, targeted at older people. At the beginning, news media were heavily used to raise awareness. Following an increased awareness, events demonstrating the benefit of long-term management of physiological conditions were held in some retirement communities, raising the want. When the want was identified, awareness spread more widely.

Figure 1. Awareness, want, and adoption gap ratio (AWAG) segment matrix.

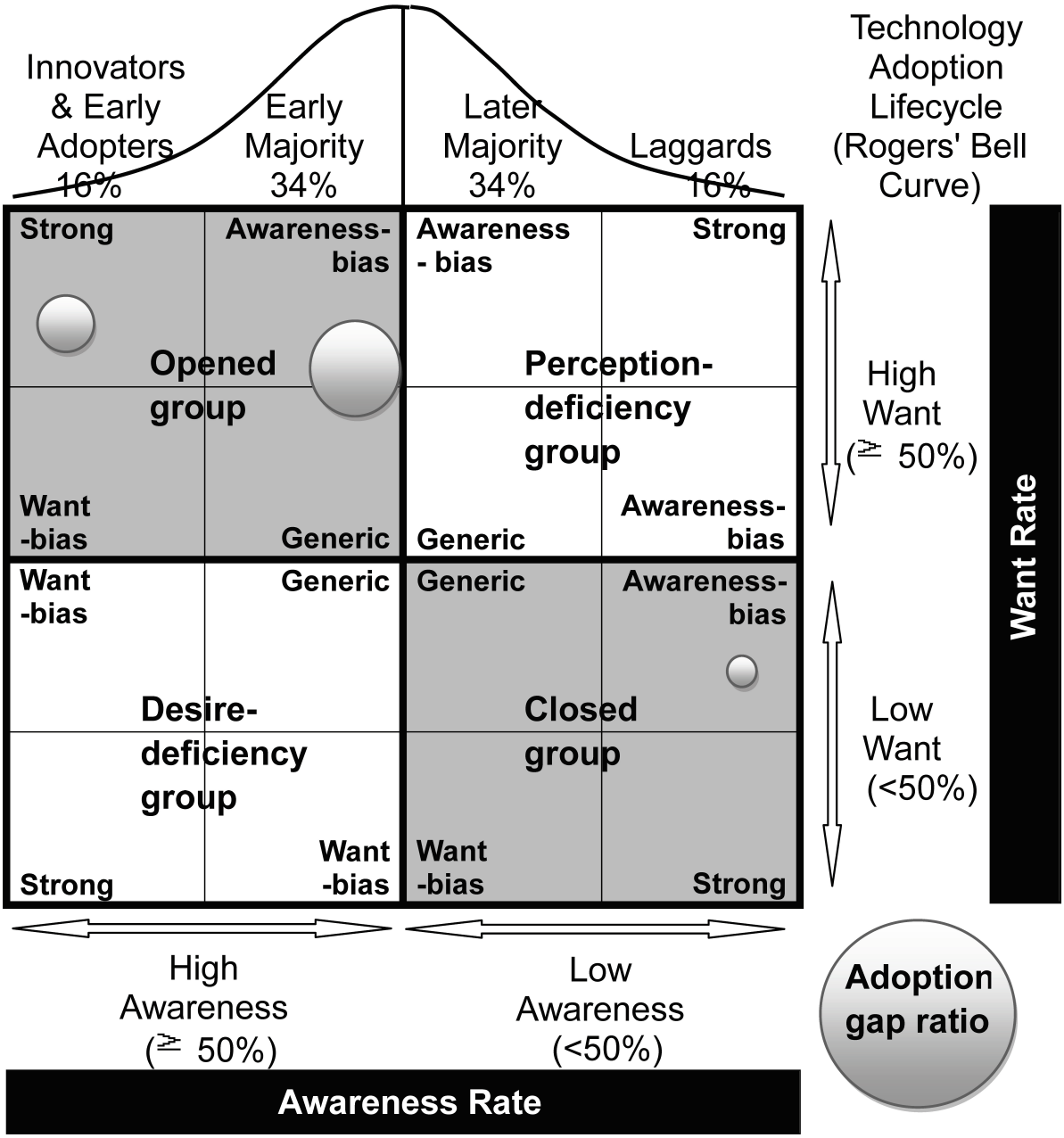
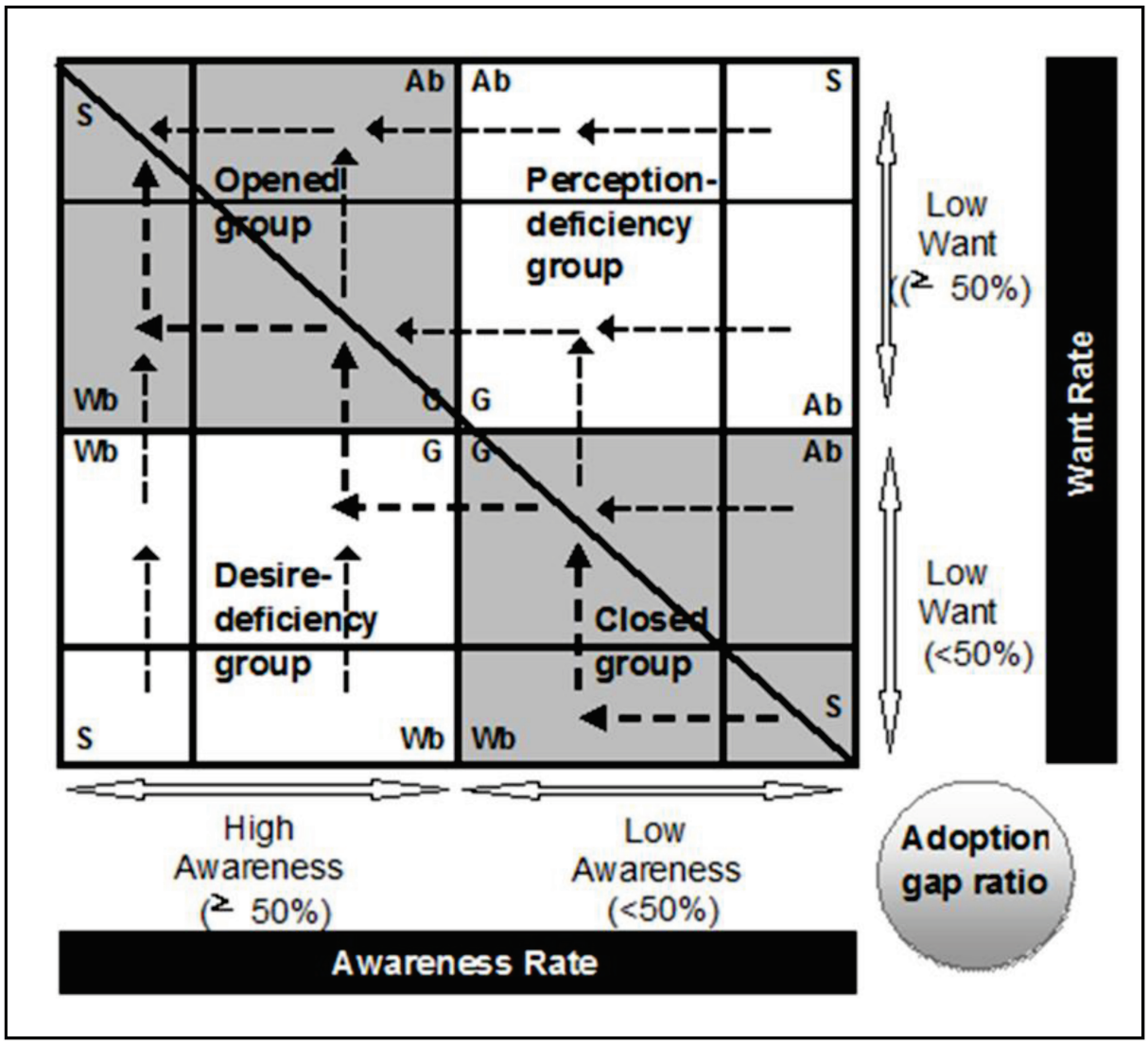


Figure 2. Improving directions for each region in the awareness, want, and adoption gap ratio (AWAG) segment matrix. Ab = awareness-bias; G = generic; S = strong; Wb = want-bias.



Adoption Gap Ratio Analysis in the AWAG Segment Matrix

Based on H4, H5, and H6, the adoption rate of a product or service is highly related to the corresponding awareness and want rates of consumers. However, even if consumers are aware of a new service or product, it does not follow that they will choose or use it. Therefore, the adoption rate is bound to the awareness and want rates. Under this situation, it is not proper to compare the adoption rates of products or services directly because they are under different levels of awareness and want (ie, the room for adoption promotion should be limited under current awareness and want). Therefore, Liang [53] proposed the adoption gap ratio analysis to explore the gap between adoption and awareness or want. The adoption gap ratio ($g(x)$gx) for service x is defined as shown in Figure 3.

The adoption gap ratio is the proportion of the adoption rate for a product or service thatcan be promoted under the current awareness or want rates. The range of the adoption gap ratio is from 0% to 100%. Among those who are already aware of or in want of an eHealth service, the adoption gate rate represents the percentage of people who have never used the service. The gate rate is 0% when the adoption rate is equal to the minimum value of the awareness and want rates. The gate rate is close to 100% when almost no one currently uses the product or service. Using the proposed adoption gap ratio, we can thus evaluate the effectiveness of adoption promotion more accurately. Several studies have found that perceived ease of use, perceived usefulness, and self-efficacy have direct effects on user attitude [60,61]. Therefore, when the gate rate is large, additional management and promotion strategies, such as enhancing the user friendliness of product or service functions or promoting adoption by education, should be used.

Figure 3. Equation for calculating the adoption gap ratio for service x.

$$g(x) = \frac{min\left(Pr\left(A(x)\right), Pr\left(W(x)\right)\right) - Pr\left(U(x)\right)}{min\left(Pr\left(A(x)\right), Pr\left(W(x)\right)\right)} \times 100\%$$

$$g(x) = \frac{min\left(Pr\left(A(x)\right), Pr\left(W(x)\right)\right) - Pr\left(U(x)\right)}{min\left(Pr\left(A(x)\right), Pr\left(W(x)\right)\right)} \times 100\%$$

where *Pr(A(x))* represents awareness rate for eHealth service*x*;
Pr(W(x)) represents the want rate for eHealth service*x*; and
Pr(U(x)) Pr(U(x)) represents the adoption rate of eHealth service*x*.

Empirical Subjects

In 2002, 62.6% of hospitals in Taiwan had developed their own websites. Most of these hospitals agreed that applying Internet technology could improve service quality and work efficiency and that the Internet would have a huge influence on the delivery of medical websites [62]. In 2011, online reservation, electronic medical records, online inquiry for medical treatment, and online drug information services were already offered by all first-tier teaching hospitals and medical centers and by most second-tier teaching hospitals. The digital medical service (DMS) is now popular in Taiwan.

According to statistics compiled by the Ministry of the Interior in Taiwan, there were 1,490,801 elderly people in December 1996, representing 7.10% of Taiwan's total population. This figure met the criteria of an old-age society set by the United Nations [63]. The elderly population has increased since then. In fact, it accounted for up to 10.21% of the population in 2010. As its population is aging fast, Taiwan has to cope with problems resulting from the continuing increase in the number of old people who need care. Therefore, the Department of Industrial Technology of the Ministry of Economic Affairs in Taiwan started planning and implementing a flagship project for technological innovation of health care service in 2006; the digital home care service (DHCS) is one of its major promotional services [64]. This project aims to bring essential care and benefits to elderly people, enabling them to live their lives with well-being, safety, convenience, and respect. Given that the DHCS has been promoted for only 5 years in Taiwan, the service has not yet gained in popularity.

To compare the digital divide on awareness, demand, and adoption of eHealth services in the different technological life cycles, the DMS and DHCS were selected as empirical subjects.

Survey Method and Questionnaire

A telephone survey was conducted to evaluate the awareness of, want for, and use of DMS and DHCS. The survey method and questionnaire are presented in Multimedia Appendix 1.

Results

Sample Structure

In all, 3074 Taiwanese respondents aged 15 years and older were interviewed by telephone. The demographic, computer ownership, and Internet access profile of the respondents is shown in Table 2. The sample distributions of gender, age, and geographic area are as homogeneous as the population distribution ($P > .05$). Based on the survey, 90.66% of the respondents had computers at home and 70.07% had Internet access from anywhere.

Table 2. Profile of respondents: Demographic variables, computer ownership, and Internet access

Demographics	n	%
Total	3074	100.0
Gender		
Male	1518	49.38
Female	1556	50.62
Age range (years)		
15−24	509	16.6
25−34	621	20.2
35−44	568	18.5
45−54	565	18.4
55−64	367	11.9
≥65	444	14.4
Education level		
Below primary school	384	12.5
Junior high school	249	8.1
Senior high school	1014	32.99
Junior college	434	14.1
University	847	27.6
Graduate and above	146	4.8
Marital status		
Single	1065	34.65
Married or cohabiting	1876	61.03
Other[a]	133	4.3
Geographic area		
Northern Area	974	31.7
Central Area	753	24.5
Southern Area	663	21.6
Eastern Area	87	3
Taipei City	365	11.9
Kaohsiung City	232	7.5
Personal monthly income (US $)		
≤450	651	21.2
451–650	211	6.9
651–950	501	16.3
951–1250	433	14.1
1251–1550	279	9.1
1551–2250	219	7.1
≥2251	140	4.6
Don’t know/no answer	641	20.8
Family monthly income (US $)		
No income or unstable	104	3.4
≤650	156	5.1

Demographics	n	%
651–1250	478	15.5
1251–1850	561	18.3
1851–2450	445	14.5
2451–3050	316	10.3
3051–4650	295	9.6
≥4651	155	5.0
Don’t know/no answer	563	18.3
Computer ownership or Internet access		
Computer ownership		
Yes	2787	90.66
No	286	9.3
Internet access		
Yes	2154	70.07
No	919	29.9

[a] Divorced, separated, or widowed.

Digital Divide Across the Demographic Variables

The awareness, want, and adoption rates of each demographic group for DMS and DHCS are shown in Table 3 and Table 4, respectively. There was a digital divide in awareness, want, and adoption in DMS across all the demographic variables, excluding gender ($P < .05$, see Multimedia Appendix 2: Table 5). Between females and males, there was no digital divide in awareness and want, whereas there was a digital divide in adoption. For DMS, the adoption rate among females is higher than that among males.

Table 3. Rates of awareness, want, and adoption by demographics and the corresponding regions in the awareness, want, and adoption gap ratio (AWAG) segment matrix for the digital medical service

			Stage							
			Awareness		Want		Adoption		Adoption gap ratio	Region in awareness–want segment matrix[a]
Variable	ID[b]	Item	n	%	n	%	n	%	%	
Total	T		2667	86.8	2390	77.8	1370	44.6	42.7	O_Wb
Gender	G1	Male	1323	87.1	1179	77.7	649	42.8	44.9	O_Wb
	G2	Female	1344	86.4	1211	77.9	721	46.3	40.5	O_Wb
Age range (years)	A1	15−24	402	78.9	399	78.4	189	37.0	52.8	O_G
	A2	25−34	579	93.2	524	84.4	343	55.2	34.5	O_Wb
	A3	35−44	539	94.9	470	82.8	282	49.6	40.1	O_Wb
	A4	45−54	510	90.3	454	80.3	245	43.3	46.0	O_Wb
	A5	55−64	321	87.5	280	76.4	143	39.0	49.0	O_Wb
	A6	≥65	315	71.1	263	59.3	169	38.0	35.9	O_G
Education level	E1	Below primary school	257	66.9	221	57.7	124	32.4	43.8	O_G
	E2	Junior high school	187	75.1	174	69.7	72	29.0	58.4	O_G
	E3	Senior high school	886	87.5	764	75.4	385	38.0	49.6	O_Wb
	E4	Junior college	410	94.5	372	85.9	223	51.4	40.1	O_S
	E5	University	784	92.6	720	84.9	469	55.3	34.9	O_Wb
	E6	Graduate and above	142	97.2	139	95.2	97	66.1	30.6	O_S
Marital status	M1	Single	918	86.2	863	81.1	466	43.8	46.0	O_Wb
	M2	Married or cohabiting	1685	89.9	1473	78.5	886	47.2	39.8	O_Wb
	M3	Other[c]	63	47.2	54	40.7	17	13.1	68.0	C_G
Geographic area	L1	Northern Area	868	89.2	775	79.6	442	45.4	43.0	O_Wb
	L2	Central Area	646	85.7	573	76.1	310	41.1	46.0	O_Wb
	L3	Southern Area	568	85.6	515	77.7	305	46.1	40.7	O_Wb
	L4	Eastern Area	76	86.9	74	84.6	37	41.9	50.5	O_Wb
	L5	Taipei City	341	93.5	304	83.4	194	53.3	36.1	O_Wb
	L6	Kaohsiung City	168	72.4	148	64.0	82	35.4	44.7	O_G
Personal monthly income (US $)	P1	≤450	520	79.9	474	72.9	265	40.7	44.2	O_G
	P2	451–650	181	85.7	161	76.3	90	42.9	43.7	O_Wb
	P3	651–950	460	91.9	414	82.7	240	47.8	42.2	O_Wb
	P4	951–1250	413	95.4	369	85.1	219	50.5	40.7	O_S
	P5	1251–1550	261	93.6	237	85.2	149	53.5	37.2	O_S
	P6	1551–2250	216	98.6	192	88.0	150	68.6	22.0	O_S
	P7	≥2251	137	97.8	124	88.5	62	44.4	49.9	O_S
	P8	Don’t know/no answer	479	74.8	419	65.3	195	30.4	53.5	O_G

Variable	ID[b]	Item	Stage						Adoption gap ratio	Region in awareness–want segment matrix[a]
			Awareness		Want		Adoption			
			n	%	n	%	n	%	%	
Family monthly income (US $)	F1	No income or unstable	88	84.9	70	67.1	39	37.8	43.7	O_G
	F2	≤650	120	76.9	106	67.8	53	33.8	50.1	O_G
	F3	651–1250	398	83.2	347	72.6	178	37.3	48.6	O_G
	F4	1251–1850	501	89.2	462	82.3	257	45.8	44.3	O_Wb
	F5	1851–2450	412	92.6	379	85.1	239	53.6	37.0	O_S
	F6	2451–3050	296	93.5	274	86.9	176	55.6	36.0	O_S
	F7	3051–4650	280	94.8	253	85.9	164	55.8	35.1	O_S
	F8	≥4651	147	94.8	135	86.9	89	57.5	33.8	O_S
	F9	Don’t know/no answer	425	75.5	364	64.6	174	30.9	52.2	O_G
Computer ownership	C1	Yes	2496	89.6	2259	81.0	1305	46.8	42.2	O_Wb
	C2	No	170	59.5	132	46.0	65	22.7	50.8	D_G
Internet access	I1	Yes	1965	91.2	1816	84.3	1114	51.7	38.6	O_Wb
	I2	No	701	76.3	574	62.5	255	27.8	55.5	O_G

[a] Groups are opened (O), desire-deficiency (D), perception-deficiency (P), and closed (C); regions are strong (S), generic (G), awareness-bias (Ab), and want-bias (Wb).

[b] Item identifier.

[c] Divorced, separated, or widowed.

Table 4. Rates of awareness, want, and adoption by demographics and the corresponding regions in the awareness, want, and adoption gap ratio (AWAG) segment matrix for the digital home care service

			Stage							
			Awareness		Want		Adoption		Adoption gap ratio	Region in awareness–want segment matrix[a]
Variable	ID[b]	Item	n	%	n	%	n	%	%	
Total	T		1563	50.9	2139	69.6	150	4.9	93.0	O_G
Gender	G1	Male	806	53.1	1055	69.5	84	5.5	92.1	O_G
	G2	Female	757	48.7	1084	69.7	66	4.3	93.9	P_G
Age range (years)	A1	15−24	250	49.1	366	71.9	31	6.1	91.5	P_G
	A2	25−34	318	51.2	440	70.8	32	5.1	92.8	O_G
	A3	35−44	315	55.5	415	73.1	23	4.0	94.5	O_G
	A4	45−54	334	59.2	432	76.5	23	4.1	94.6	O_G
	A5	55−64	173	47.2	246	67.1	15	4.1	93.9	P_G
	A6	≥65	173	39.0	239	53.9	27	6.0	88.9	P_G
Education level	E1	Below primary school	120	31.3	196	51.2	15	4.0	92.3	P_G
	E2	Junior high school	114	45.7	154	61.5	8	3.1	95.0	P_G
	E3	Senior high school	520	51.3	711	70.1	44	4.4	93.8	O_G
	E4	Junior college	262	60.4	330	76.1	20	4.5	94.0	O_G
	E5	University	451	53.3	636	75.1	55	6.4	91.4	O_G
	E6	Graduate and above	96	65.8	113	77.5	8	5.7	92.6	O_G
Marital status	M1	Single	541	50.8	765	71.8	62	5.8	92.0	O_G
	M2	Married or cohabiting	989	52.7	1325	70.6	86	4.6	93.5	O_G
	M3	Other[c]	34	25.1	49	37.0	2	1.7	95.3	C_G
Geographic area	L1	Northern Area	502	51.5	689	70.8	53	5.4	92.4	O_G
	L2	Central Area	410	54.4	536	71.2	44	5.8	91.8	O_G
	L3	Southern Area	323	48.8	443	66.7	29	4.4	93.5	P_G
	L4	Eastern Area	46	52.2	70	80.6	3	3.4	95.7	O_G
	L5	Taipei City	192	52.6	270	74.1	10	2.6	96.4	O_G
	L6	Kaohsiung City	91	39.4	131	56.5	12	5.1	91.0	P_G
Personal monthly income (US $)	P1	<450	277	42.6	422	64.8	33	5.1	92.1	P_G
	P2	451–650	117	55.6	146	69.0	13	6.2	91.0	O_G
	P3	651–950	258	51.5	358	71.4	24	4.8	93.3	O_G
	P4	951–1250	228	52.7	332	76.7	23	5.4	93.0	O_G
	P5	1251–1550	174	62.6	206	74.0	15	5.3	92.9	O_G
	P6	1551–2250	135	61.5	178	81.2	14	6.2	92.4	O_G
	P7	≥2251	93	66.6	112	79.6	12	8.7	89.0	O_G
	P8	Don’t know/no answer	280	43.7	387	60.4	16	2.5	95.8	P_G

Variable	ID[b]	Item	Stage						Adoption gap ratio	Region in awareness–want segment matrix[a]
			Awareness		Want		Adoption			
			n	%	n	%	n	%	%	
Family monthly income (US $)	F1	No income or unstable	40	38.3	60	57.7	1	0.6	99.0	P_G
	F2	≤650	71	45.3	92	58.7	5	3.2	94.6	P_G
	F3	651–1250	210	44.0	330	69.2	24	5.0	92.8	P_G
	F4	1251–1850	320	57.0	416	74.1	26	4.7	93.7	O_G
	F5	1851–2450	246	55.3	325	73.0	29	6.4	91.2	O_G
	F6	2451–3050	171	54.2	237	74.9	17	5.5	92.7	O_G
	F7	3051–4650	182	61.8	232	78.7	22	7.6	90.4	O_G
	F8	≥4,651	87	56.2	113	72.8	12	8.0	89.1	O_G
	F9	Don’t know/no answer	236	41.9	334	59.3	14	2.4	95.9	P_G
Computer ownership	C1	Yes	1473	52.9	2026	72.7	143	5.1	92.9	O_G
	C2	No	90	31.4	114	39.7	6	2.3	94.3	C_G
Internet access	I1	Yes	1192	55.3	1615	75.0	125	5.8	92.3	O_G
	I2	No	371	40.4	524	57.0	25	2.7	95.3	P_G

[a] Groups are opened (O), desire-deficiency (D), perception-deficiency (P), and closed (C); regions are strong (S), generic (G), awareness-bias (Ab), and want-bias (Wb).

[b] Item identifier.

[c] Divorced, separated, or widowed.

With respect to the digital divide in DHCS, except for gender, the *P* values of the chi-square independent tests for awareness and want are all less than .05 (see Multimedia Appendix 2: Table 6), indicating that there was a digital divide in awareness and want in DHCS across all the demographic variables, excluding gender. Between females and males, there was no digital divide in want, whereas there was a digital divide in awareness. The awareness rate of males was higher than that of females. The above results support hypotheses H1, H2, and H3.

Adoption Rate is Always Bound to Awareness and Want Rates

In the paired proportion test between adoption and awareness or want rates for the demographic groups, there are significant differences between adoption and awareness or want rates across the demographic groups ($P < .05$, see Multimedia Appendix 2: Table 7 and Table 8). Given that all the adoption rates are less than the awareness and want rates (Table 3, Table 4), the results support hypotheses H4 and H5.

Want Rate is Not Necessarily Bound to Awareness Rate

In the paired proportion test between awareness and want rates for the demographic groups, there are significant differences between awareness and want rates for most demographic groups ($P < .05$, see Multimedia Appendix 2: Table 7 and Table 8). However, the awareness rates are significantly greater than the want rates for DMS, which is an existing eHealth service in Taiwan. The want rates are significantly greater than the awareness rates for DHCS, which is a new eHealth service in Taiwan (Table 3, Table 4). This indicates that the want rate is not necessarily dependent on the awareness rate. Thus, hypotheses H6, H6-1, and H6-2 are supported by the above results.

Conditional Relationship among Awareness, Want, and Adoption Rates

In the independent proportion test of want rates between unawareness and awareness for DMS and DHCS, there are significant differences for most demographic groups ($P < .05$, see Multimedia Appendix 2: Table 9 and Table 10). Furthermore, for all the demographic groups, want rates given unawareness are less than want rates given awareness. Thus, H7 is supported. For most of the demographic groups, the adoption or awareness rates under “wanted” are greater than those under “unwanted” for DMS and DHCS. Thus, hypotheses H8 and H9 are mostly supported.

AWAG Segment Matrix Analysis

With respect to the analysis of the AWAG segment matrix for DMS, most of the segments, except for people without a computer or with marriage status of divorced, separated, or widowed, are positioned under the opened group. This is not surprising given that DMS was already well established in Taiwan. However, although most of the segments belong to the opened group, their degrees of openness are different. In general, individuals with high levels of personal and family monthly incomes, as well as those with education levels of graduate and above, belong to the strong opened group, whereas those who are either younger or older, have low education and family

monthly income levels, are living in Kaohsiung City, or have no computer or Internet access belong to the generic opened group. Those without a computer belong to the generic desire-deficiency group with relatively low want, indicating that they may have high awareness of DMS, but their needs may not be identified. People with a marriage status of divorced, separated, or widowed belong to the generic closed group. With relatively low DMS awareness, they may be encountering some barriers in obtaining information (Figure 4, Figure 5, Figure 6, Figure 7, Figure 8, Figure 9, Figure 10, Figure 11, Table 3).

The adoption gap ratios of segments for DMS range from 22.0% to 68.0%. All segments still have room to promote adoption. Among people with a marriage status of Other, 68% of those who were aware or in want of DMS did not adopt the service. The adoption gap ratios are near or above 50% among people aged between 25 and 34 or between 55 and 64 years; having education levels of junior high school or senior high school; having marriage status of Other; living in the eastern area; having personal monthly incomes above US $2251 and family monthly incomes less than 1250, or did not know or refused to answer questions on personal or family monthly income; and had no computer or Internet access. These findings indicate a huge potential area where DMS adoption can be promoted.

With respect to analysis of the AWAG segment matrix for DHCS, segments are separately positioned in the opened, perception-deficiency, and closed groups. Segments of females; aged between 15 and 24, 55 and 64, and above 65 years; having education levels below primary school or junior school; living in the southern area or Kaohsiung City; having personal monthly incomes of less than US $450, or did not know or refused to answer questions on personal income; having minimal family monthly incomes that were unstable or less than US $1250, or did not know or refused to answer questions on family income; or had no Internet access are positioned in the generic perception-deficiency group. This indicates that the awareness of DHCS for those segments is relatively low. People with marital status of divorced, separated, or widowed, as well as those without a computer, belong to the generic closed group. Both their awareness and want rates are relatively low. Other segments not mentioned above all belong to the generic opened group (Figure 4, Figure 5, Figure 6, Figure 7, Figure 8, Figure 9, Figure 10, Figure 11, Table 3). The gap ratios of all segments are all greater than or near 90%. Given that DHCS is a new eHealth service in Taiwan, there are huge potential areas where DHCS adoption can be promoted for all segments.

Figure 4. Awareness, want, and adoption gap ratio (AWAG) segment matrix for the digital medical service (DMS) and digital home care service (DHCS).

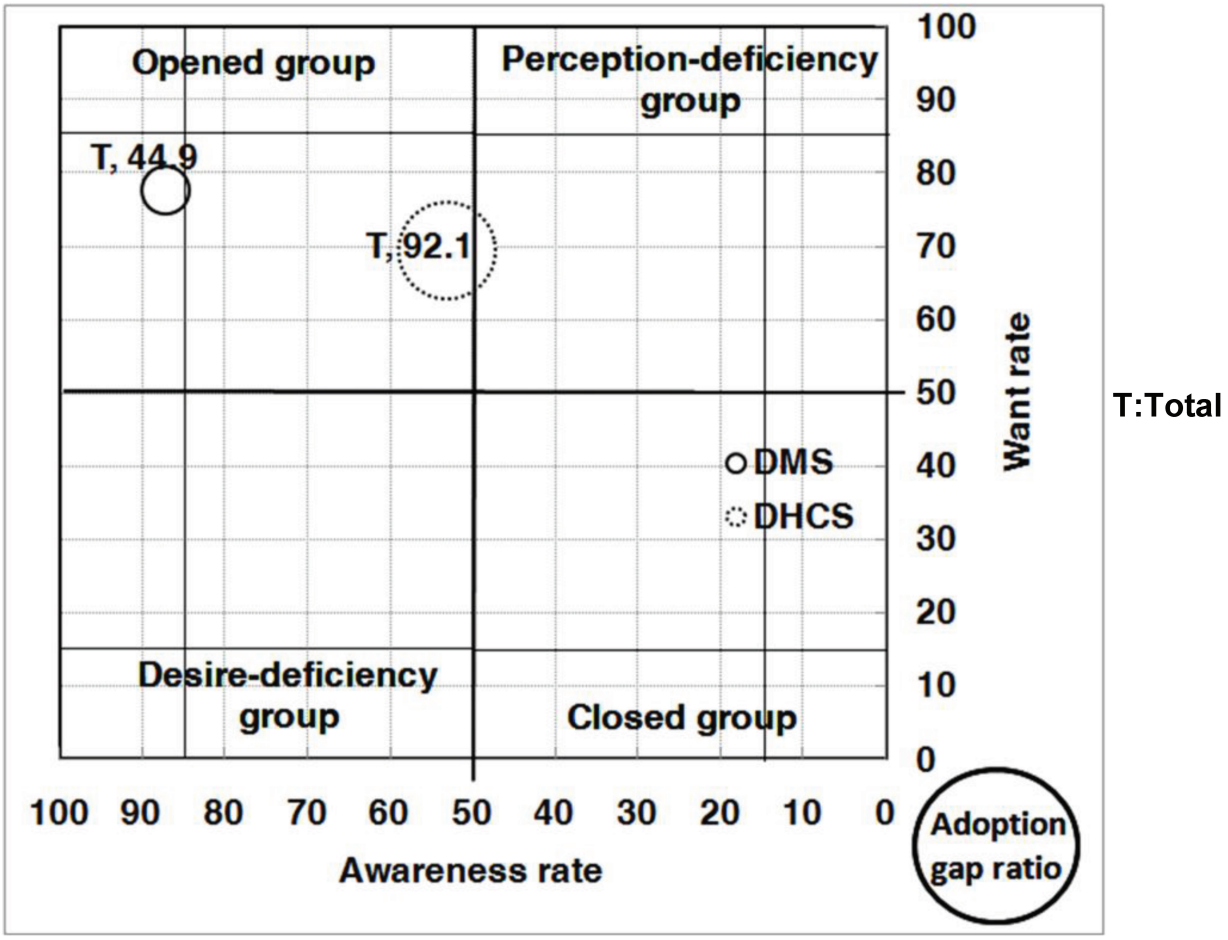


Figure 5. Awareness, want, and adoption gap ratio (AWAG) segment matrix by age (A; years) for the digital medical service (DMS) and digital home care service (DHCS).

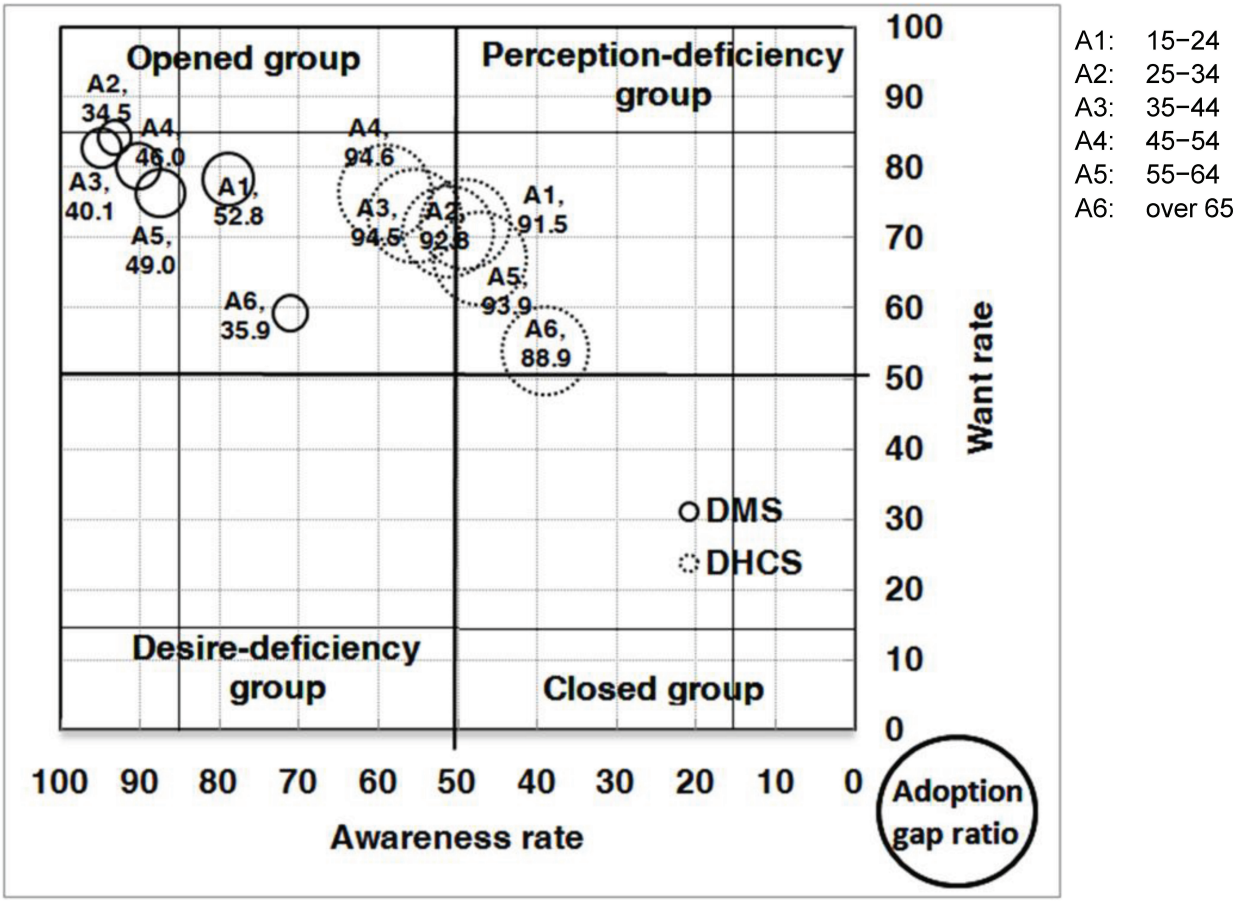


Figure 6. Awareness, want, and adoption gap ratio (AWAG) segment matrix by educational level (E) for the digital medical service (DMS) and digital home care service (DHCS).

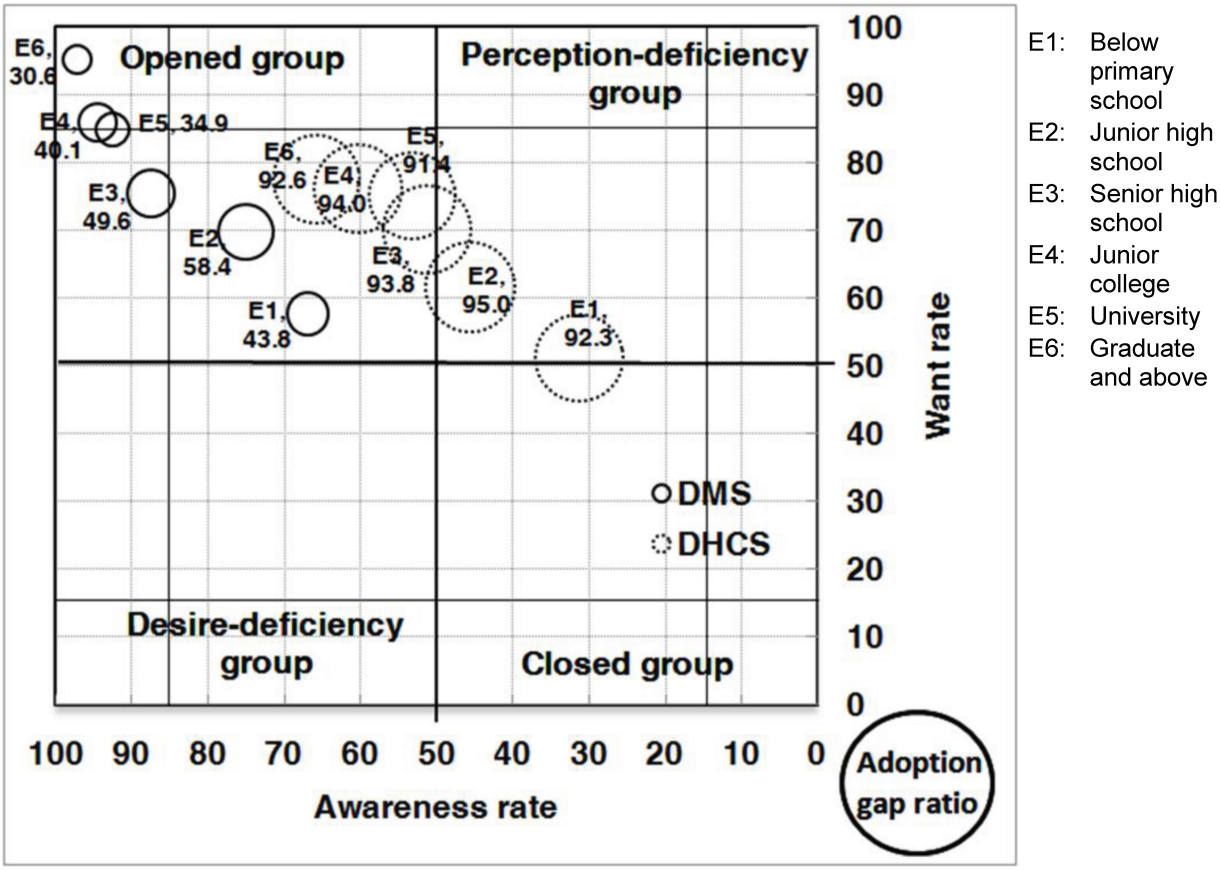


Figure 7. Awareness, want, and adoption gap ratio (AWAG) segment matrix by marital status (M) for the digital medical service (DMS) and digital home care service (DHCS).

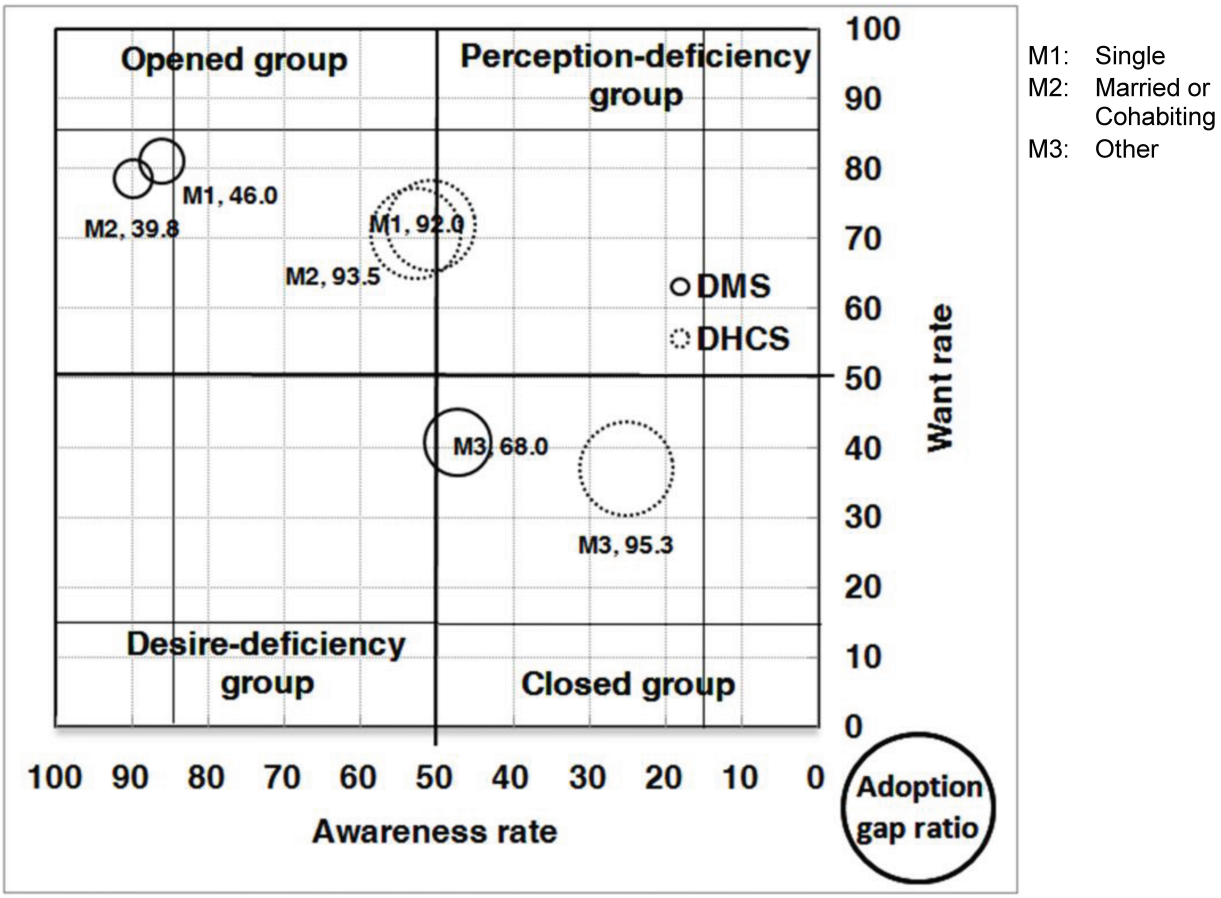


Figure 8. Awareness, want, and adoption gap ratio (AWAG) segment matrix by geographic area (L) for the digital medical service (DMS) and digital home care service (DHCS).

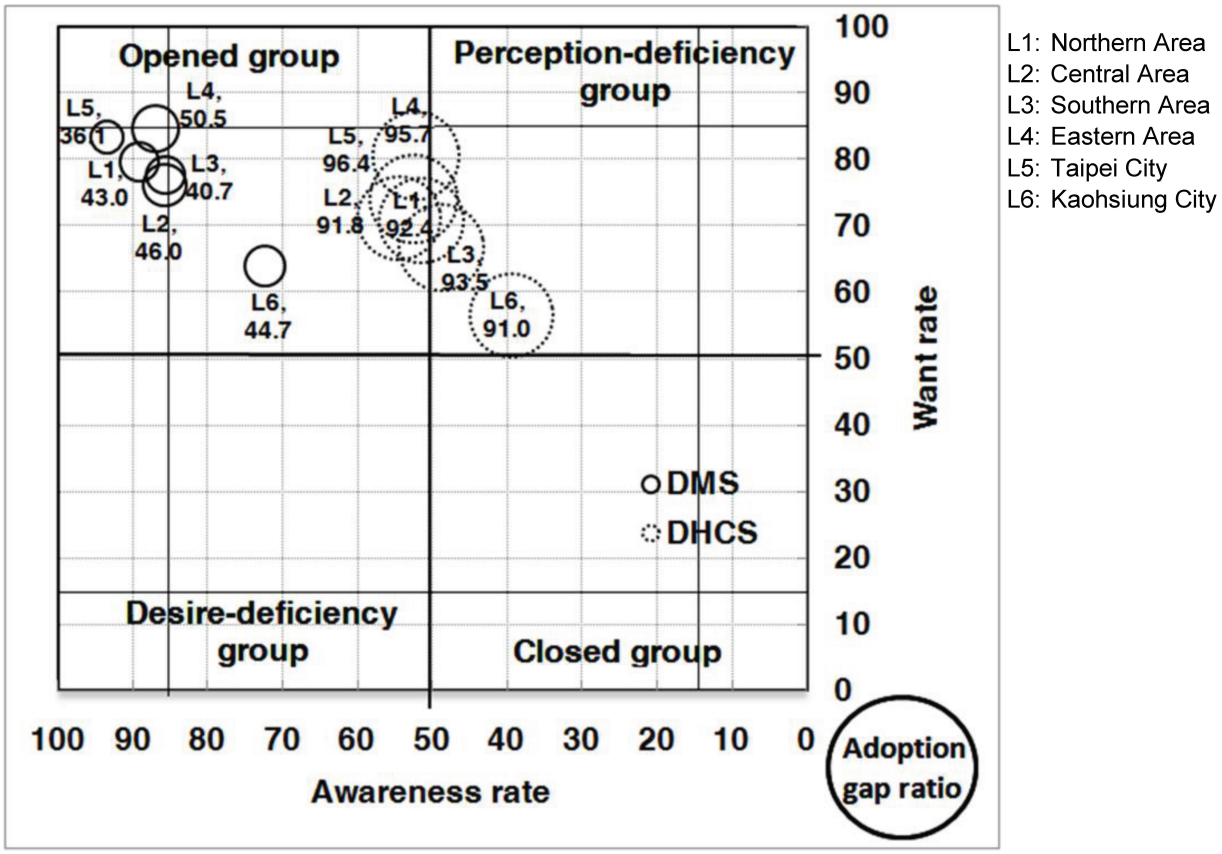


Figure 9. Awareness, want, and adoption gap ratio (AWAG) segment matrix by personal monthly income (P) for the digital medical service (DMS) and digital home care service (DHCS).

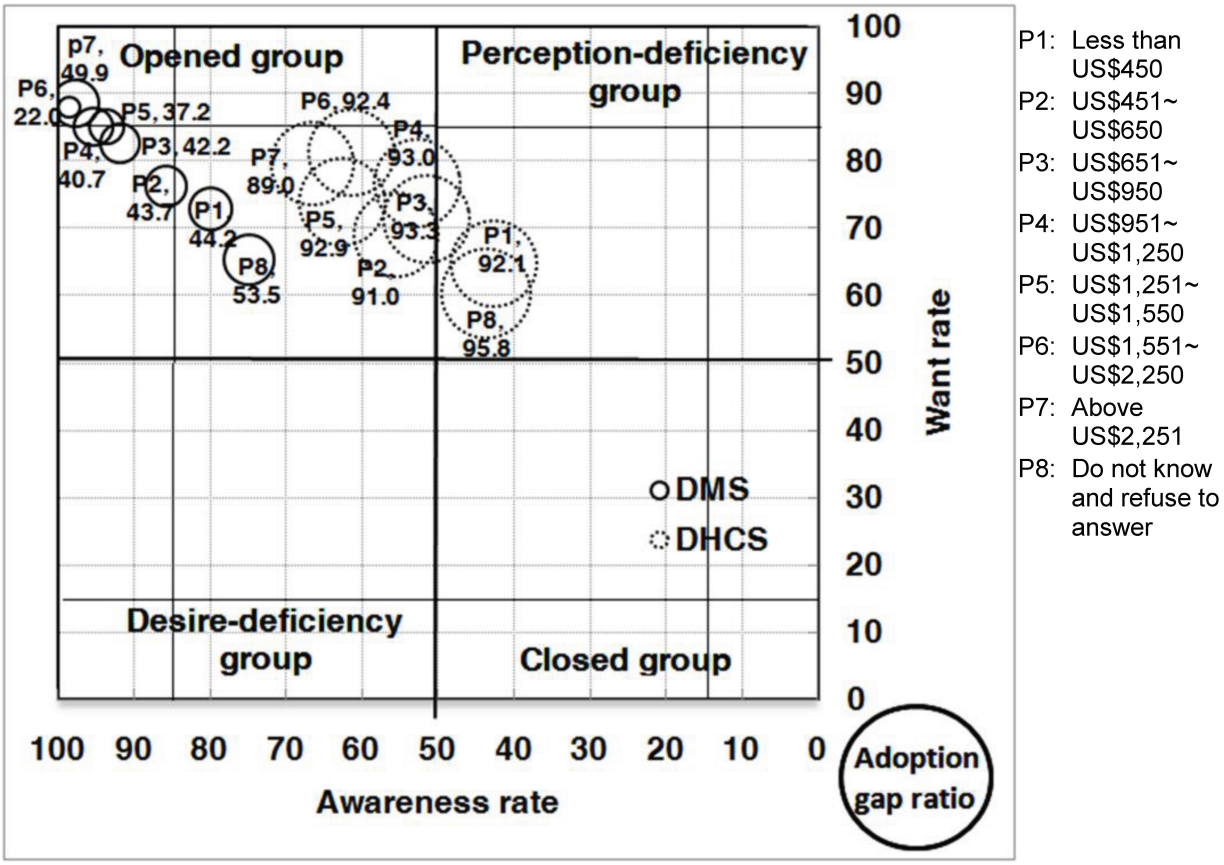


Figure 10. Awareness, want, and adoption gap ratio (AWAG) segment matrix by family monthly income (F) for the digital medical service (DMS) and digital home care service (DHCS). DK = don't know; RA = refused to answer.

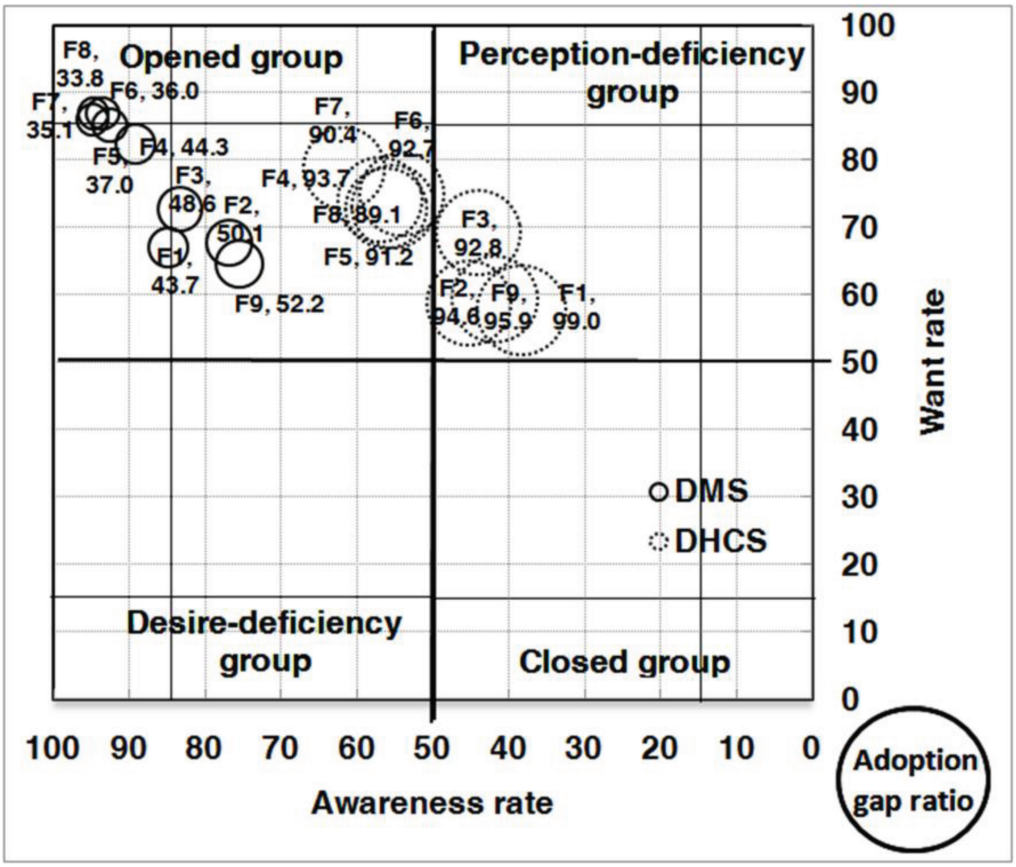
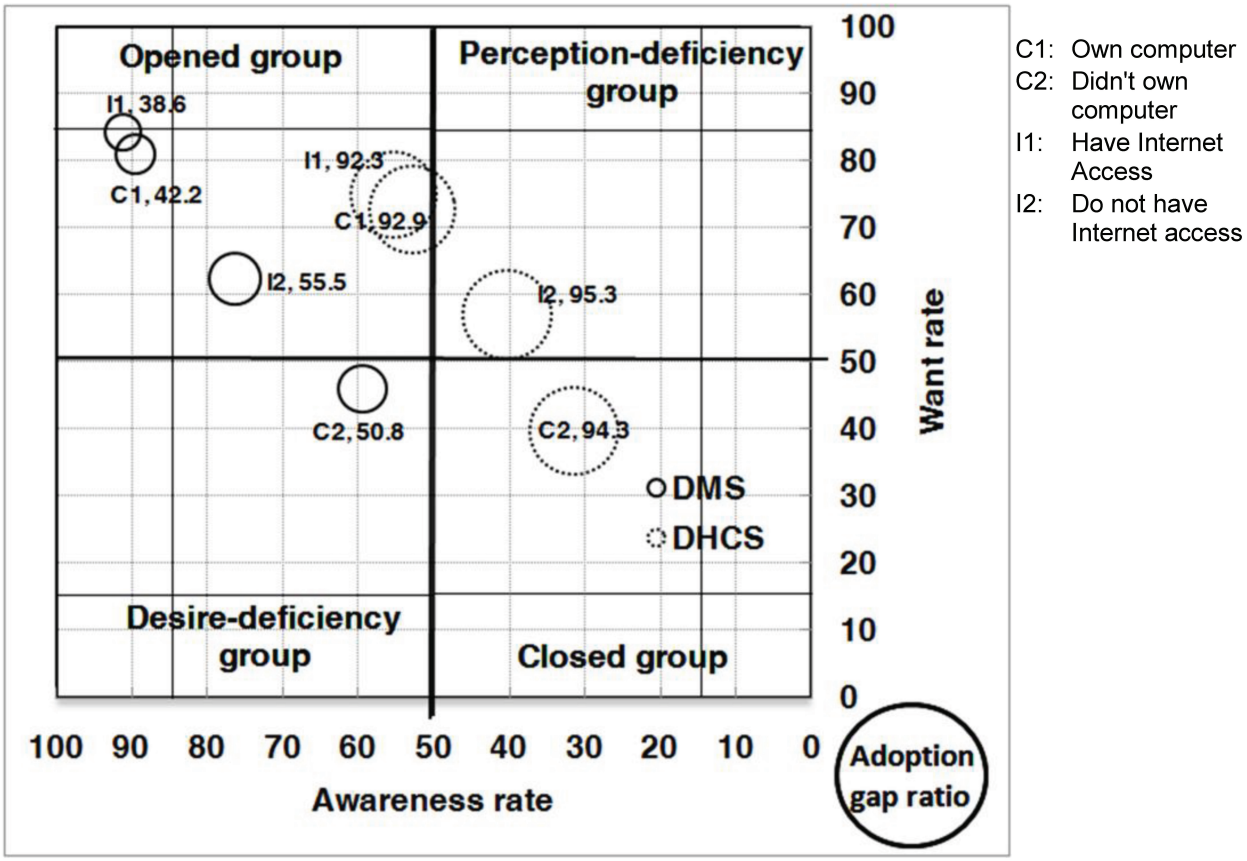


Figure 11. Awareness, want, and adoption gap ratio (AWAG) segment matrix by computer ownership (C) and Internet access (I) for the digital medical service (DMS) and digital home care service (DHCS).



Discussion

The results of this study show that digital divides in DMS and DHCS exist across certain demographic variables. In addition, the study has proven that the want rate is not always bound to the awareness rate. The want rate is usually bound to the awareness rate for existing services, such as DMS. However, DHCS is an innovative e-service in Taiwan; thus, for this service, the want rate is higher than the awareness rate. This study has also proven that awareness and want have reciprocal effects. Adoption may be pulled with rising awareness and want. A higher awareness rate may result in a higher want rate, and people in need of some eHealth services have a higher awareness rate than those who are not in need of the service. Therefore, the innovation diffusion process should start from awareness, followed by want, and awareness and want will gradually rise through the cycle of technology push and consumer pull, and pull adoption.

Using AWAG segment matrix analysis led to several conclusions. With respect to DMS, most segments belong to the opened group. Based on the adoption gap ratio analysis, all the gap values are higher than 22%, signifying that there is room for raising strategies for DMS adoption.

For DMS, segments with high levels of personal and family monthly incomes, as well as segments with education levels of graduate and above, all belong to the strong opened group and are the primary target markets. The marketing strategy of “hold” and the action of “keep up the good work” are suggested. Compared with other segments, the adoption gap ratio for those with personal monthly incomes above US $2251 is the highest and is near 50%. This segment should be ranked first in terms of adoption promotion strategies.

Segments with members who are either younger or older, have low education levels, have low family monthly income levels, live in Kaohsiung City, and have no Internet access belong to the generic opened group for DMS. For these, the “hold and improve strategy” and the action of “keep up the good work and keep raising the awareness and want” are recommended. These segments constitute the third target market for DMS. Other segments, except for people without a computer and with marriage status of Other, constitute the secondary target market for DMS. Thus, the “hold and improve strategy” with the action of “keep up the good work and keep raising the awareness” is suggested.

People without computers belong to the generic desire-deficiency group and are nontarget markets for DMS. The marketing strategy of “improve (create)” and the action of “keep creating the want” are suggested. The adoption gap ratio for this segment is 50.8%; in this segment, half of those who do not adopt DMS are in want of DMS. Thus, the adoption promotion strategy should also be used at once. People with marriage status of Other belong to the generic closed group. The adoption gap ratio for this group is the highest; in this

segment, 68% of those who do not adopt DMS are in want of DMS. Thus, the "evaluate then improve (raise)" strategy and the actions of "evaluate the potential of the segment then keep raising the awareness or want" and "promote the adoption" are suggested.

With respect to DHCS, half of the segments belong to the opened group, and one-third of the segments belong to the perception-deficiency group. According to the adoption gap ratio analysis, because DHCS is a new eHealth service in Taiwan, all the gap values are higher than or near 90%, indicating that there is a huge room for raising DHCS adoption.

Segments with members who are female, young, late-middle aged, or elderly; have low education levels; live in the southern area or Kaohsiung City; have low or unstable personal or family incomes or refuse to answer questions on income; or without Internet access belong to the generic perception-deficiency group for DHCS. These segments are potential target markets for DHCS, and the marketing strategy "improve (spread)" and the action of "keep spreading the awareness" are suggested.

People who are divorced, separated, or widowed, or without computers belong to the generic closed group for DHCS. These two segments are nontarget markets for DHCS, and the "evaluate then leave or evaluate then improve (raise)" strategy and the action of "evaluate the potential of the segment then choose an action between "maintain status quo" and "keep raising the awareness or want" should be used. Other segments for DHCS all belong to the generic opened group. These constitute the third target market for DHCS. The "hold and improve strategy" and the action of "keep up the good work" and "keep raising the awareness and want" are thus suggested.

Conclusion

This study has proposed the AWAG segment matrix analysis and analyzed the digital divides in DMS and DHCS across different demographic groups. From the results of this study, the digital divide in awareness and want across different demographic groups can be easily observed by cross-segmenting the awareness and want rates. Marketing strategies have also been clearly established. The adoption gap ratio between adoption and awareness or want rates is large for DMS and even larger for DHCS. These indicate that adoption does not closely follow peoples' awareness or want, and that a huge digital divide in adoption exists in DHS and DHCS. Adoption education and promotion programs should therefore be used.

For marketing managers in business, government, or other related institutions, the AWAG segment matrix provides a simple and clear method for analyzing the digital divides and differentiating between target and nontarget markets. Moreover, it helps managers adjust their market strategies and allocate their budget more effectively from an objective and customer-oriented viewpoint.

Suggestions for Future Study

For further research, the AWAG segment matrix can be revised by adding "satisfaction with eHealth service" into the analysis. The AWAG segment matrix can also be extended to analyze differential concerns on information security among the different segments mentioned in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey methodology and questionnaire.

[PDF File (Adobe PDF File), 194KB - jmir_v14i1e11_app1.pdf]

Multimedia Appendix 2

Tables 5-11.

[PDF File (Adobe PDF File), 319KB - jmir_v14i1e11_app2.pdf]

References

1. DeLuca JM, Enmark R. E-health: the changing model of healthcare. Front Health Serv Manage 2000;17(1):3-15. [Medline: 11184427]
2. Kwankam SY. What e-Health can offer. Bull World Health Organ 2004 Oct;82(10):800-802. [Medline: 15643805]
3. Internet World Stats. Miniwatts Marketing Groups. 2011 Dec 20. Internet Usage Statistics: The Internet Big Picture: World Internet Users and Population Stats URL: http://www.internetworldstats.com/stats.htm [accessed 2012-01-08] [WebCite Cache ID 64XETFIwQ]
4. Sharp C. Advancing the promise of e-health through information technology. Online 2010;34(2):34-39.
5. Harrison JP, Lee A. The role of e-Health in the changng health care environment. Nurs Econ 2006;24(6):283-8, 279; quiz 289. [Medline: 17266004]
6. Haider M, Kreps GL. Forty years of diffusion of innovations: utility and value in public health. J Health Commun 2004;9 Suppl 1:3-11. [doi: 10.1080/10810730490271430] [Medline: 14960400]

7. Preston J, Brown FW, Hartley B. Using telemedicine to improve health care in distant areas. Hosp Community Psychiatry 1992 Jan;43(1):25-32. [Medline: 1544643]
8. Brown FW. Rural telepsychiatry. Psychiatr Serv 1998 Jul;49(7):963-964. [Medline: 9661235]
9. Hilty DM, Servis ME, Nesbitt TS, Hales RE. The use of telemedicine to provide consultation-liaison service to the primary care setting. Psychiatr Ann 1999;29:421-427.
10. Hilty DM, Yellowlees PM, Cobb HC, Bourgeois JA, Neufeld JD, Nesbitt TS. Models of telepsychiatric consultation--liaison service to rural primary care. Psychosomatics 2006;47(2):152-157 [FREE Full text] [doi: 10.1176/appi.psy.47.2.152] [Medline: 16508028]
11. Straker N, Mostyn P, Marshall C. The use 0f two-way TV in bringing mental health services to the inner city. Am J Psychiatry 1976 Oct;133(10):1202-1205. [Medline: 970494]
12. Nesbitt TS, Hilty DM, Kuenneth CA, Siefkin A. Development of a telemedicine program: a review of 1,000 videoconferencing consultations. West J Med 2000 Sep;173(3):169-174. [Medline: 10986176]
13. Ida T, Asai S. The regional cost gap of the Japanese local telecommunications service. Pap Reg Sci 2005;84(2):183-199. [doi: 10.1111/j.1435-5957.2005.00012.x]
14. Zhang ML, Wolff RS. Crossing the digital divide: cost-effective broadband wireless access for rural and remote areas. IEEE Commun Mag 2004;42(2):99-105. [doi: 10.1109/MCOM.2003.1267107]
15. Gray NJ, Klein JD, Cantrill JA, Noyce PR. Adolescent girls' use of the Internet for health information: issues beyond access. J Med Syst 2002 Dec;26(6):545-553. [Medline: 12385536]
16. 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] [Medline: 19615467]
17. Wilson EV, Lankton NK. Modeling patients' acceptance of provider-delivered e-health. J Am Med Inform Assoc 2004;11(4):241-248 [FREE Full text] [doi: 10.1197/jamia.M1475] [Medline: 15064290]
18. Fesenmaier J. Laboratory for Community and Economic Development, University Of Illinois at Urbana-Champaign. 2003 Jan 9. Executive Summary: Measuring Outcomes of Digital Divide Investment to Community Technology Centers URL: http://www.communitydevelopment.uiuc.edu/PDFs/ExecutiveSummary.pdf [accessed 2012-01-08] [WebCite Cache ID 64XEkJqjA]
19. Katz JE, Rice RE. Social Consequences of Internet Use: Access, Involvement, and Interaction.. Cambridge, MA: MIT Press; 2002.
20. Fung V, Ortiz E, Huang J, Fireman B, Miller R, Selby JV, et al. Early experiences with e-health services (1999-2002): promise, reality, and implications. Med Care 2006 May;44(5):491-496. [doi: 10.1097/01.mlr.0000207917.28988.dd] [Medline: 16641669]
21. Prieger JE, Hu WM. The broadband digital divide and the nexus of race, competition, and quality. Inform Econ Policy 2008;20(2):150-167. [doi: 10.1016/j.infoecopol.2008.01.001]
22. Liff S, Shepherd A. An Evolving Gender Digital Divide? Internet Issue Brief no 2. Oxford: Oxford Internet Institute; 2004. URL: http://educ.ubc.ca/faculty/bryson/565/genderdigdiv.pdf [accessed 2012-01-08] [WebCite Cache ID 64XF1ilGN]
23. Aqii SV, Moghaddam AI. Bridging the digital divide: the role of librarians and information professionals in the third millennium. Electronic Library 2008;26(2):226-237. [doi: 10.1108/02640470810864118]
24. Roy S, Ghose S. Internet adoption as a two-stage transition: converting internet non-users to internet users and to online buyers. Int J Market Res 2006;48(3):321-349 [FREE Full text]
25. Livingstone S, Helsper E. Balancing opportunities and risks in teenagers' use of the Internet: the role of online skills and internet self-efficacy. New Media Soc 2009;12(2):309-329. [doi: 10.1177/1461444809342697]
26. Rice RE, Katz JE. Comparing internet and mobile phone usage: digital divides of usage, adoption, and dropouts. Telecomm Policy 2003;27(8-9):597-623. [doi: 10.1016/S0308-5961(03)00068-5]
27. Fairlie R, London R, Pastor M, Rosner R. Center for Justice, Tolerance & Community, University of California Santa Cruz. 2004. Race, Youth, & the Digital Divide: A Research and Policy Initiative URL: http://cjtc.ucsc.edu/PowerPoints/Offline_01_unlinked_brownbag_mp.pdf [accessed 2012-01-08] [WebCite Cache ID 64XFToqdj]
28. Samal L, Hutton HE, Erbelding 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] [Medline: 19941085]
29. Grubesic TH. Inequities in the broadband revolution. Ann Reg Sci 2003;37(2):263-289. [doi: 10.1007/s001680300123]
30. Dey A, Reid B, Godding R, Campbell A. Perceptions and behaviour of access of the Internet: a study of women attending a breast screening service in Sydney, Australia. Int J Med Inform 2008 Jan;77(1):24-32. [doi: 10.1016/j.ijmedinf.2006.12.002] [Medline: 17218149]
31. Prieger JE. The supply side of the digital divide: is there equal availability in the broadband Internet access market? Econ Inq 2003;41(2):346-363. [doi: 10.1093/ei/cbg013]
32. LaRose R, Gregg JL, Strover S, Straubhaar J, Carpenter S. Closing the rural broadband gap: promoting adoption of the Internet in rural America. Telecomm Policy 2007;31(6-7):359-373. [doi: 10.1016/j.telpol.2007.04.004]
33. Billon M, Marco R, Lera-Lopez F. Disparities in ICT adoption: a multidimensional approach to study the cross-country digital divide. Telecomm Policy 2009;33(10-11):596-610. [doi: 10.1016/j.telpol.2009.08.006]

34. McTavish FM, Pingree S, Hawkins R, Gustafson D. Cultural differences in use of an electronic discussion group. J Health Psychol 2003 Jan;8(1):105-117. [doi: 10.1177/1359105303008001447] [Medline: 22113904]
35. Ginossar T, Nelson S. Reducing the health and digital divides: a model for using community-based participatory research approach to e-health interventions in low-income Hispanic communities. J Comput Med Commun 2010;15(4):530-551. [doi: 10.1111/j.1083-6101.2009.01513.x]
36. Swartz SH, Cowan TM, Batista IA. Using claims data to examine patients using practice-based Internet communication: is there a clinical digital divide? J Med Internet Res 2004 Jan 9;6(1):e1 [FREE Full text] [doi: 10.2196/jmir.6.1.e1] [Medline: 15111267]
37. Ko CH, Yen JY, Chen CC, Chen SH, Yen CF. Gender differences and related factors affecting online gaming addiction among Taiwanese adolescents. J Nerv Ment Dis 2005 Apr;193(4):273-277. [Medline: 15805824]
38. Ko CH, Yen JY, Yen CF, Lin HC, Yang MJ. Factors predictive for incidence and remission of internet addiction in young adolescents: a prospective study. Cyberpsychol Behav 2007 Aug;10(4):545-551. [doi: 10.1089/cpb.2007.9992] [Medline: 17711363]
39. Escoffery C, Diiorio C, Yeager KA, McCarty F, Robinson E, Reisinger E, et al. Use of computers and the Internet for health information by patients with epilepsy. Epilepsy Behav 2008 Jan;12(1):109-114. [doi: 10.1016/j.yebeh.2007.07.013] [Medline: 17959420]
40. Wright TP. Factors affecting the cost of airplanes. J Aeronaut Sci 1936;3:122-128.
41. Hellman K, Burst A. The Customer Learning Curve: Creating Profits From Marketing Chaos. Mason, OH: American Marketing Association/Thomson South-western; 2004.
42. Rogers EM. Diffusion of Innovations, 3rd edition. New York, NY: Free Press; 1983.
43. Prins R, Verhoef PC. Marketing communication drivers of adoption timing of a new e-service among existing customers. J Marketing 2007;71(2):169-183. [doi: 10.1509/jmkg.71.2.169]
44. Steenkamp J-BEM, Gielens K. Consumer and market drivers of the trial probability of new consumer packaged goods. J Consum Res 2003;30(3):368-384. [doi: 10.1086/378615]
45. McGuire WJ. An information-processing model of advertising effectiveness. In: Davis HL, Silk AJ, editors. Behavioral and Management Science in Marketing. New York, NY: Ronald Press; 1978:156-180.
46. Lavidge RC, Steiner GA. A model for predictive measurements of advertising effectiveness. J Marketing 1961;25:59-62.
47. Bruyn AD, Lilien GL. A multi-stage model of word-of-mouth influence through viral marketing. Int J Res Marketing 2008;25(3):151-163. [doi: 10.1016/j.ijresmar.2008.03.004]
48. Hauser JR, Urban GL. A normative methodology for modeling consumer response to innovation. Oper Res 1977;25:579-619. [doi: 10.1287/opre.25.4.579]
49. Barry TE, Howard DJ. A review and critique of the hierarchy of effects in advertising. Int J Advert 1990;9(2):121-135.
50. Eng TY, Quaia G. Strategies for improving new product adoption in uncertain environments: a selective review of the literature. Ind Marketing Manag 2009;38(3):275-282. [doi: 10.1016/j.indmarman.2008.01.003]
51. Ottum BD, Moore WL. The role of market information in new product success/failure. J Prod Innov Manag 1997;14(4):258-273. [doi: 10.1111/1540-5885.1440258]
52. Rosen C, Dorsett D, Jack J. A proline-rich region in the Zeste protein essential for transvection and white repression by Zeste. Genetics 1998 Apr;148(4):1865-1874 [FREE Full text] [Medline: 9560400]
53. Liang TH. The awareness and want matrix with adoption gap ratio analysis for e-service diffusion effect. Cyberpsychol Behav Soc Netw 2011 Mar;14(3):131-140. [doi: 10.1089/cyber.2009.0349] [Medline: 21117973]
54. Dixon DR. The behavioral side of information technology. Int J Med Inform 1999 Dec;56(1-3):117-123. [Medline: 10659940]
55. Dowling GR. The Art and Science of Marketing: Marketing for Marketing Managers. Oxford: Oxford University Press; 2004.
56. Martin MJC. Managing Innovation and Entrepreneurship in Technology-Based Firms. New York, NY: Wiley; 1994.
57. Peter JP, Donnelly JHJ. A Preface to Marketing Management. Boston, MA: McGraw-Hill/Irwin; 2003.
58. Kotler P. Marketing Management: Analysis, Planning, Implementation and Control, 9th edition. New York, NY: Prentice-Hall; 1997.
59. Mercer D. A two-decade test of product life cycle theory. Br J Manag 1993;4(4):269-274. [doi: 10.1111/j.1467-8551.1993.tb00063.x]
60. Chen J, Yen D, Chen K. The acceptance and diffusion of the innovative smart phone use: a case study of a delivery service company in logistics. Inform Manag 2009;46(4):241-248. [doi: 10.1016/j.im.2009.03.001]
61. Scott JE, Walczak S. Cognitive engagement with a multimedia ERP training tool: assessing computer self-efficacy and technology acceptance. Inform Manag 2009;46(4):221-232. [doi: 10.1016/j.im.2008.10.003]
62. Kuo NC, Kuo PC, Yang MC. A study on building hospitals' websites and related factors. Taiwan J Public Health 2002;21(6):439-448.
63. Chan HS, Lin CC. Taiwan National Policy Foundation. 2005 Mar 17. Crisis and Challenge of Taiwan's Aging Society: NPF backgrounder URL: http://old.npf.org.tw/PUBLICATION/SS/094/SS-B-094-006.htm [accessed 2012-01-08] [WebCite Cache ID 64XHEt0ng]

64. Ministry of Economic Affairs. Smart Living: 2010 White Paper on Taiwan Industrial Technology (Part II). Taipei, Taiwan: Ministry of Economic Affairs; 2010:1-10.

Abbreviations

AWAG: awareness, want, and adoption gap ratio
DHCS: digital home care service
DMS: digital medical service
ITAM: information technology adoption model

Edited by G Eysenbach; submitted 18.10.10; peer-reviewed by K Bennett; comments to author 12.04.11; revised version received 13.10.11; accepted 25.10.11; published 13.02.12.

Please cite as:
Liang TH
Matrix Analysis of the Digital Divide in eHealth Services Using Awareness, Want, and Adoption Gap
J Med Internet Res 2012;14(1):e11
URL: http://www.jmir.org/2012/1/e11/
doi:10.2196/jmir.1670
PMID:22329958

Original Paper

Acceptability and Preliminary Feasibility of an Internet/CD-ROM-Based Education and Decision Program for Early-Stage Prostate Cancer Patients: Randomized Pilot Study

Michael A Diefenbach[1], PhD; Nihal E Mohamed[2], PhD; Brian P Butz[3], PhD; Natan Bar-Chama[4], MD; Richard Stock[5], MD; Jamie Cesaretti[6], MD; Waleed Hassan[7], MD; David Samadi[4], MD; Simon J Hall[4], MD

[1]Mount Sinai School of Medicine, Departments of Urology & Oncological Sciences, Mount Sinai School of Medicine, New York, NY, United States

[2]Departments of Urology & Oncological Sciences, Mount Sinai School of Medicine, New York, NY, United States

[3]College of Engineering, Temple University, Philadelphia, PA, United States

[4]Department of Urology, Mount Sinai School of Medicine, New York, NY, United States

[5]Department of Radiation Oncology, Mount Sinai School of Medicine, New York, NY, United States

[6]Florida Radiation Oncology Group, Jacksonville, FL, United States

[7]Department of Urology, Tawam Hospital, Al Ain, United Arab Emirates

Corresponding Author:
Michael A Diefenbach, PhD
Mount Sinai School of Medicine
Departments of Urology & Oncological Sciences
Mount Sinai School of Medicine
Box 1272
One Gustave L. Levy Place
New York, NY, 10029-0751
United States
Phone: 1 212 241 0751
Fax: 1 212 876 0751
Email: michael.diefenbach@mountsinai.org

Abstract

Background: Prostate cancer is the most common cancer affecting men in the United States. Management options for localized disease exist, yet an evidence-based criterion standard for treatment still has to emerge. Although 5-year survival rates approach 98%, all treatment options carry the possibility for significant side effects, such as erectile dysfunction and urinary incontinence. It is therefore recommended that patients be actively involved in the treatment decision process. We have developed an Internet/CD-ROM-based multimedia Prostate Interactive Educational System (PIES) to enhance patients' treatment decision making. PIES virtually mirrors a health center to provide patients with information about prostate cancer and its treatment through an intuitive interface, using videos, animations, graphics, and texts.

Objectives: (1) To examine the acceptability and feasibility of the PIES intervention and to report preliminary outcomes of the program in a pilot trial among patients with a new prostate cancer diagnosis, and (2) to explore the potential impact of tailoring PIES treatment information to participants' information-seeking styles on study outcomes.

Methods: Participants (n = 72) were patients with newly diagnosed localized prostate cancer who had not made a treatment decision. Patients were randomly assigned to 3 experimental conditions: (1) control condition (providing information through standard National Cancer Institute brochures; 26%), and PIES (2) with tailoring (43%) and (3) without tailoring to a patient's information-seeking style (31%). Questionnaires were administrated before (t1) and immediately after the intervention (t2). Measurements include evaluation and acceptability of the PIES intervention, monitoring/blunting information-seeking style, psychological distress, and decision-related variables (eg, decisional confidence, feeling informed about prostate cancer and treatment, and treatment preference).

Results: The PIES program was well accepted by patients and did not interfere with the clinical routine. About 79% of eligible patients (72/91) completed the pre- and post-PIES intervention assessments. Patients in the PIES groups compared with those in the control condition were significantly more likely to report higher levels of confidence in their treatment choices, higher levels of helpfulness of the information they received in making a treatment decision, and that the information they received was

emotionally reassuring. Patients in the PIES groups compared with those in the control condition were significantly less likely to need more information about treatment options, were less anxious about their treatment choices, and thought the information they received was clear ($P < .05$). Tailoring PIES information to information-seeking style was not related to decision-making variables.

Conclusions: This pilot study confirms that the implementation of PIES within a clinical practice is feasible and acceptable to patients with a recent diagnosis of prostate cancer. PIES improved key decision-making process variables and reduced the emotional impact of a difficult medical decision.

(J Med Internet Res 2012;14(1):e6) doi:10.2196/jmir.1891

KEYWORDS

Multimedia; software; prostate cancer; patient education; treatment decision making; treatment; decision making

Introduction

Prostate cancer is the most common cancer and the second leading cause of cancer-related deaths among American men [1]. In 2011, it is estimated that approximately 217,730 men will have a prostate cancer diagnosis and approximately 32,050 men will die of the disease [1]. Early-stage prostate cancer patients (ie, who present with a tumor that is confined to the prostate, and who have no regional lymph node or distant metastasis, or T1–2N0M0 [2,3]) can choose between several treatment options—surgery (ie, prostatectomy) and radiation therapy (ie, external radiation, brachytherapy, or CyberKnife robotic radiosurgery)—or active surveillance. Although both surgical and radiotherapy approaches have excellent cancer control, each treatment option is characterized by a distinctive pattern of potentially long-lasting urinary, bowel, and sexual dysfunction [3]. In the absence of an evidence-based standard for therapy, it is important for patients to understand how different treatment options will influence their immediate and long-term quality of life. Arriving at a treatment decision can be quite challenging for patients. Most patients are highly distressed after a cancer diagnosis, yet they are required to absorb a large amount of medical information that is often presented in language fraught with medical and probabilistic terms [4-6], and they have to resolve often contradictory medical opinions from consulting physicians of different medical subspecialties. Making a treatment decision under these circumstances is difficult and may lead to emotional distress and subsequent decisional regret, especially if the chosen treatment and its side effects decrease the patients' quality of life [7].

Electronic and traditional print-based education decision-making materials have been used for patients with localized prostate cancer to enhance treatment decision making. Both types of materials have been shown to improve disease-specific knowledge and to facilitate decision making [8]. Yet few comprehensive Web-based resources are available for prostate cancer patients that combine unbiased treatment information culled from the existing literature, with physicians answering frequently asked questions and testimonials from prostate cancer survivors (eg, [9-11]). We present results from a small pilot study that examined acceptability, feasibility, and preliminary outcomes of a state-of-the-art multimedia intervention, Prostate Interactive Educational System (PIES) [11] designed to educate patients about their treatment options and to facilitate their treatment decision process.

PIES Description

PIES is designed to present disease- and treatment-relevant information through a variety of electronic media (ie, text, graphics, video clips, and animation) and self-navigational aids, and is a truly innovative, state-of-the-science multimedia preparatory aid for prostate cancer patients (see Multimedia Appendix, Multimedia Appendix 2, Multimedia Appendix 3, and Multimedia Appendix 4). Conceptually, PIES serves as a virtual health center that patients visit to obtain prostate cancer-relevant information. On entering the system, patients are greeted by a health educator who gives them an overview of PIES and its contents (ie, physicians' offices, library, and support group room). Physicians are represented by videos of actual doctors who answer questions about prostate cancer treatment within their area of specialization. The library contains books about treatment options and side effects illustrated with graphics, photographs, and animations. The support group room allows patients to listen to the experiences of prostate cancer survivors who have undergone treatment. Groups of three survivors stratified by treatment type are represented through videos, and patients have the option to learn how these survivors have coped with the decision-making process, posttreatment issues, and potential side effects. In addition, the emotional aspects of a prostate cancer diagnosis have been addressed throughout the software. The library books include statements that attempt to normalize the diagnosis, to reduce negative affect and discourage avoidant coping, and to encourage problem-solving coping—for example, "talk about your feelings and concerns, and learn as much as you can about your cancer and treatment." In the support group room, survivors talk about how they have coped with the disease, treatment decision making, and side effects. Additionally, physicians provide normalizing statements while talking about treatment and side effects (eg, "You are not alone in this diagnosis, thousands of men are being diagnosed with prostate cancer every year"). The detailed developmental and usability testing process has been described elsewhere [11].

Theoretical Basis of PIES

Self-regulation theory [12,13] guided the development of PIES and the selection of study outcomes. Self-regulation theory postulates two parallel processing arms, one for cognitive and one for affective representations of a health threat or stimulus

[12,13]. Cognitive processes are characterized by illness representation attributes such as knowledge or beliefs about a threat or stimulus, its causation, consequence, duration, controllability, and overall understanding (ie, illness cohesion). At the same time, affective processes occur in reaction to the stimulus. The importance of affect in decision-making processes has recently been recognized [14]. Negative affect triggered by the cognitive appraisal of a health threat could in turn influence further information processing and bias decision-making processes [14]. Thus, both cognitive and affective factors influence information processing, decision making, and ultimately behavior, and therefore need to be addressed when providing information to patients.

Research has shown that information processing preferences such as high- versus low-monitoring information-seeking styles also play a role in information processing and decision making [15,16]. A monitoring information-seeking style can be conceptualized as a tendency for individuals to select, encode, interpret, react affectively to, and manage threatening medical health information in either a high- or low-monitoring information-seeking style. A *high-monitoring* style is characterized by an increased need for information and by scanning for and magnifying stress-related cues relevant to one's health, whereas a *low-monitoring* style is characterized by a reduced need for information, distraction, and minimization of health cues. Previous studies found that monitoring was significantly associated with differential cognitive–affective responses and coping with health-related stressors (eg, a cancer diagnosis) [17-19]. Thus, we explored the influence of text information offered through PIES' library tailored to a high- or low-information style on information processing and decision variables. We expected that tailored information would improve information processing and facilitate decision making.

The purpose of this study was twofold: (1) to examine the acceptability and feasibility of the PIES intervention and to report preliminary outcomes of the program in a randomized pilot trial among patients with newly diagnosed prostate cancer, and (2) to explore the potential impact of tailoring PIES messages to participants' information-seeking styles on study outcomes. Following recommendations of assessing feasibility in applied intervention research [18], we evaluated the feasibility of the PIES intervention based on (1) *acceptability of PIES* (ie, successful implementation of the PIES intervention; relevance and acceptability), (2) *recruitment and retention* (ie, participation and attrition rates), (3) *timeline* (ie, ability to offer the intervention as planned shortly after diagnosis and to assess study outcomes), and (4) *preliminary outcomes* (ie, within the context of a small pilot study, preliminary evidence for the effect of the program on decision-making variables [16]). We hypothesized that patients assigned to use PIES (1) would evaluate the PIES program positively as indicated by high usability ratings, high satisfaction with the information presented, and high ratings of the helpfulness of the program in making a treatment decision, and (2) would be more satisfied with the information they received, be better informed, display less decisional distress, and have more decisional certainty than the men who were assigned to the control condition (ie, received National Cancer Institute [NCI] brochures on prostate cancer); and that (3) men in the tailored PIES group would be more satisfied with the information they received, be better informed, display less decisional distress, and have more decisional certainty than the men in the PIES without tailoring group.

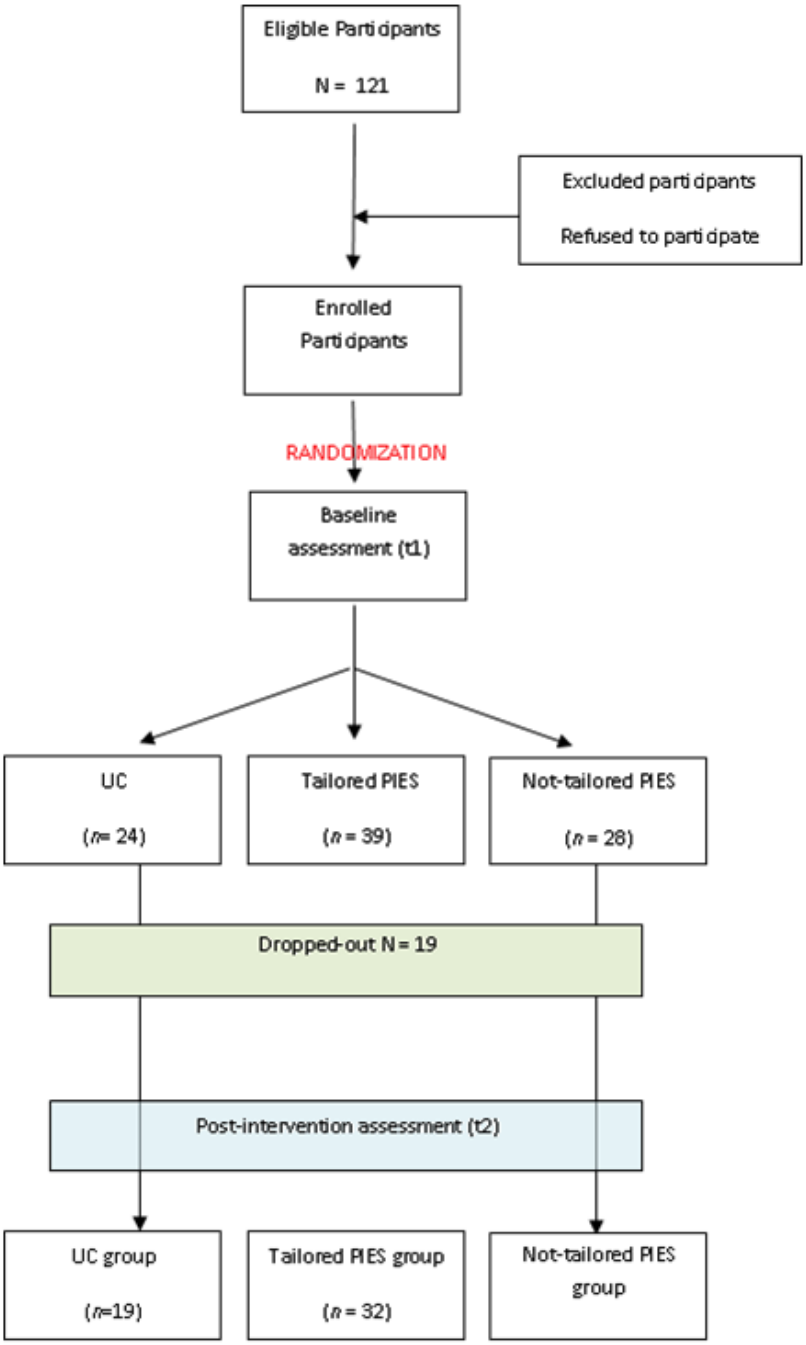
Methods

Procedure and Inclusion/Exclusion Criteria

We recruited patients with newly diagnosed prostate cancer (T1–2N0M0) who presented themselves at an urban medical center in the Northeast of the United States (n = 121) for a consultation about treatment options between June 2005 and December 2007(see Figure 1). Eligibility criteria were a diagnosis of localized prostate cancer during the past 4 to 6 weeks, ability and willingness to provide written informed consent, and fluency in English. Exclusion criteria were serious comorbidities that would limit patients' treatment options (eg, cardiovascular diseases that would prohibit a surgical approach). The institutional review board of Mount Sinai School of Medicine reviewed and approved the study.

Collaborating physicians introduced the study to eligible patients and obtained permissions for study personnel to contact patients. If patients agreed to be contacted they were telephoned by the research staff, the study was discussed in detail, and patients were asked to arrive 1 hour prior to their physician appointment. All participants provided written informed consent. To maximize the limited subject pool we randomly assigned patients following a 2:2:1 ratio into 1 of 3 study groups (PIES with tailoring, PIES without tailoring, and the control condition). All data were self-reported and collected through self-administered questionnaires with assistance by the research assistant if necessary (t1 = baseline prior to intervention/comparison conditions; t2 immediately following the intervention/comparison conditions). For tailoring purposes patients in the PIES condition also completed an abbreviated electronic version of the Monitoring/Blunting Style Scale [19]. Based on published cut-offs for monitoring/blunting style the software assigned the patient into 1 of the 2 monitoring/blunting conditions [19].

Figure 1. Participants' flow through the study. PIES = Prostate Interactive Educational System; UC = control group.



Tailoring

The PIES expert system presented all written information available in the library according to the patient's preferred monitoring style. Specifically, patients in the high-monitoring group received detailed and lengthy descriptions of disease- and treatment-related processes, enhanced by animations and graphics. Descriptions for low monitors, in contrast, were brief, lacking extensive detail, and access to animations and graphics was optional.

Comparison Condition

We designed the comparison condition as an attention control condition that incorporated elements of usual care. Patients received NCI-published brochures that are routinely provided by physicians: (1) *Understanding Treatment Choices for Prostate Cancer* [20], and (2) *What You Need to Know about Prostate Cancer* [21]. Both booklets explain basic facts about prostate cancer, and its treatment and side effects. All patients were asked to read the brochures for 45 minutes, the identical time patients in the intervention condition had to explore PIES.

Study Measures

Demographic and Clinical Characteristics

The baseline questionnaire included demographic (eg, age, ethnicity, marital status, educational, and employment) information. Self-reported medical variables were verified through a chart review (eg, prostate-specific antigen [PSA] level, time of diagnosis, and treatment preferences).

Usability Measures

Guided by previous intervention trials [22] and the recommendations of the Science Panel on Interactive Communication and Health [23], we designed and used 17 items to assess the usability of PIES and the NCI brochures with respect to improving understanding about prostate cancer treatment and its side effects, enhancing treatment decision making, and addressing concerns about the disease, treatment side effects, and cure. Examples of items used are "How useful was the information you received?" "How satisfied are you with the information you received?" "How helpful was the information you received in making a treatment decision?" "Was the information you received confusing?" Responses were endorsed on a 5-point scale (ie, 1 = not al all, to 5 = very much), with higher scores indicating higher levels of satisfaction with the material received.

Additional items to evaluate the usability of PIES were ease of use, clarity, understandability, and helpfulness of the different modules of PIES (eg, library, glossary, and visual materials), resulting in a total of 9 questions. Patients in the 2 PIES groups were asked to endorse on a 5-point scale (1 = not at all; 5 = very much) whether *information was clearly presented* and *easy to understand* (see Table 2 for more examples). Higher scores on these scales represent a more positive evaluation.

Treatment Decision Measures

We measured decisional variables adopted from the decisional conflict scale [7] with 5 items assessing treatment preferences or decision ("What is your treatment choice?"), confidence in treatment decision ("How confident are you about your treatment choice?"), feeling informed about prostate cancer ("How informed do you feel about prostate cancer?"), need for more information about treatment and side effects ("Would you prefer to have more information to make your treatment decision?"), and need for more time ("Would you prefer to have more time to make your treatment decision?"). Item responses ranged from not at all (1) to very much (5), with higher scores indicating higher levels of decisional confidence, feeling better informed, and an increased need for more time and information. Need for more time and information, confidence in treatment decision, and feeling informed about prostate cancer were measured at postintervention assessment (t2; see Table 4). Treatment preferences were measured at both baseline and postintervention assessments.

Affect Measure

We used the Impact of Event Scale-Revised (IES-R) [24] to examine psychological distress at baseline (t1). The IES-R is composed of 2 subscales that characterize 2 forms of psychological distress: intrusion (7 items; Cronbach alpha = .84) and avoidance (8 items; Cronbach alpha = .80). Items are measured on a 4-point Likert scale (0–5): not at all (0), rarely (1), sometimes (3), and often (5). The 2 subscales were highly intercorrelated (r = .68, P < .001). Accordingly, we used an overall mean score to indicate the level of psychological distress, with higher scores indicating higher levels of subjective distress (Cronbach alpha = .88).

The Monitoring/Blunting Style Scale

We used 8 items of the brief Monitoring/Blunting Style Scale [19] to assess patients' information-seeking style. The scale consists of 2 scenarios (eg, threat of job loss, going to the dentist), which are followed by 8 potential responses. Of the 8 items of each scenario, 4 describe information seeking or monitoring (eg, "I would talk to my fellow workers to see if they knew anything about the supervisor's evaluation of me"), and the remaining 4 items describe blunting responses (eg, "I would push all thoughts of being dismissed out of my mind"). A composite score is calculated by subtracting the blunting sum score from the monitoring sum score. High monitoring/low blunting scores (ie, positive scores) indicate a higher tendency toward a monitoring information-seeking style [19]. This measure was administered at baseline assessment (t1) and implemented in PIES for tailoring purposes.

Statistical Analysis

We analyzed the data with SPSS for Windows, version 16.0 (IBM Corporation, Somers, NY, USA) using descriptive statistics, t test, chi-square test, and analysis of variance procedures. Evidence for the acceptability and feasibility of the PIES intervention is demonstrated by (1) high participation and low attrition rates, (2) positive evaluation of PIES, and (3) significant differences between the 2 PIES groups and the comparison control group in postintervention outcomes. Evidence for a significant effect of tailoring the PIES intervention to the patients' monitoring/blunting styles is demonstrated by significant differences in post-PIES decision-related outcomes between the 2 PIES groups (ie, tailored versus nontailored intervention groups).

Results

Presentation of results is divided into 4 parts, which describe (1) enrollment and attrition analyses, (2) sample demographic and clinical characteristics, (3) program evaluation, and (4) the impact of PIES on study outcomes.

Enrollment and Attrition Analyses

Of the 121 referred eligible patients, 91 agreed to participate in the study and completed the baseline questionnaire (t1; 75% acceptance rate). Reasons for nonparticipation were a lack of time and interest, or having made a treatment decision.

To examine any potential bias introduced through selective attrition, we compared patients who completed the study (n = 72) with patients who did not (n = 19) on demographic (eg, age, marital status, race, employment, and education levels), clinical (eg, age, PSA level, treatment preferences), and psychological variables (eg, baseline distress). Almost 80% of patients also completed the t2 assessments (72/91, 79%), providing evidence that the PIES program can be integrated into the clinical counseling routine. The most common reason for not providing the immediate postintervention assessments was lack of time, as patients were called into the doctor's office for their appointment (ie, "running out of time" and "having a doctor appointment"). Results showed no significant differences between patients who dropped out and patients who completed the postintervention assessment (ie, all P > .05).

Sample Demographic and Clinical Characteristics

The majority of the sample (n = 91) were white (59%), were married (72%), and had a college or higher education (56%); 49% were currently employed. Average PSA level at diagnosis was 7.84 (SD 7.71) µg/L. At baseline patients expressed treatment preferences for brachytherapy (38%), prostatectomy (20%), 3-dimensional conformal radiation therapy (11%), active surveillance (13%), and other treatments (19%, eg, hormone therapy, or complementary and alternative medicine). The majority of patients (87%) had access to a home computer and the Internet (88%); additionally 29% had access to a computer at work. Examining differences between the 3 study groups with regard to baseline demographic, clinical, psychological distress variables, and monitoring/blunting information-seeking style showed no significant differences (see Table 1). Therefore, demographic and clinical variables were not included in comparative analyses as covariates. Subsequent comparative analyses included data only from patients who completed the baseline and post-PIES or control condition assessments.

Table 1. Characteristics and decisional outcomes at baseline (t1), by study group

	Total sample (n = 72)	Comparison groups			F/χ^2	*P* value
		Control condition (n = 19, 26%)	PIES[a] with tailoring (n = 32, 45%)	PIES without tailoring (n = 21, 29%)		
Demographic and clinical characteristics						
Age (years), mean (SD)	61.93 (8.08)	64.16 (8.35)	60.03 (7.77)	62.81 (8.01)	1.76	.18
PSA[b] level (µg/L), mean (SD)	7.84 (7.71)	7.69 (3.08)	7.63 (9.45)	8.25 (7.14)	0.08	.09
Time since diagnosis (weeks), mean (SD)	8.59 (18.88)	9.27 (16.79)	9.32 (123.20)	7.25 (13.58)	10.0	.98
≤High school, n (%)	30 (44%)	7 (41%)	13 (42%)	10 (50%)	4.50	.60
≥College, n (%)	38 (56%)	10 (59%)	18 (58%)	10 (50%)		
Married/with partner, n (%)	51 (72%)	15 (83%)	22 (69%)	14 (67%)	8.40	.39
Single/widowed/divorced, n (%)	20 (18%)	3 (17%)	10 (31%)	7 (33%)		
White, n (%)	41 (59%)	10 (56%)	17 (55%)	14 (66%)	5.65	.46
African American, n (%)	19 (27%)	6 (33%)	8 (26%)	5 (24%)		
Hispanic/Asian/other, n (%)	10 (14%)	2 (11%)	6 (19%)	2 (10%)		
Employed, n (%)	34 (49%)	8 (42%)	20 (63%)	6 (29%)	7.26	.12
Not employed/retired, n (%)	35 (51%)	11 (58%)	10 (37%)	14 (71%)		
Baseline treatment preferences						
Surgery, n (%)	13 (20%)	0 (0%)	8 (29%)	5 26%)	8.4	.40
External beam radiation therapy, n (%)	7 (11%)	3 (18%)	3 (11%)	1 (5%)		
Brachytherapy, n (%)	24 (38%)	7 (41%)	9 (32%)	8 (42%)		
Active surveillance, n (%)	8 (13%)	3 (18%)	2 (7%)	3 (16%)		
Other, n (%)	12 (19%)	4 (24%)	6 (21%)	2 (11%		
Baseline psychological covariates						
Psychological distress score					0.54	.58
Mean (SD)	1.70 (0.68)	1.88 (1.07)	1.68 (1.03)	1.54 (0.81)		
Range	0–4	0–4	0–4	0–3		
Monitoring style score					2.09	.13
Mean (SD)	4.27 (2.06)	4.89 (2.18)	3.75 (1.19)	4.50 (2.22)		
Range	1–8	0–8	1–8	0–8		
Blunting style score					0.05	.94
Mean (SD)	2.18 (1.59)	2.11 (1.44)	2.25 (1.86)	2.15 (1.30)		
Range	0–8	0–6	0–8	0–5		

[a] Prostate Interactive Educational System.

[b] Prostate-specific antigen.

Evaluation of PIES and Print Materials

Patients in both PIES groups and the control group were satisfied with the educational materials they received, and reported that the materials improved their understanding of (1) prostate cancer and its diagnosis, (2) treatment options and side effects (3) follow-up health care and (4) support groups and clinical trials (see Table 2). However, when comparing the 2 groups we found that patients in the control group were significantly more likely to report that the information they received (1) was confusing, (2) was too voluminous, and (3) made them more anxious about their treatment decisions (all $P < .05$; see Table 2). These patients were also significantly less likely to report that the information they received helped them make a treatment decision or was emotionally reassuring.

Table 2. Evaluation of program and print materials at postintervention assessment (t2), by study group

Usability assessment items	Total sample (n = 72)	Study group comparisons				
		Control condition (n = 19, 26%)	PIES[a] with tailoring (n = 32, 45%)	PIES without tailoring (n = 21, 29%)	F_2	*P* value
Enhancing treatment information, mean (SD)						
Provided information was useful	4.15 (0.09	4.11 (1.05)	4.25 (0.67)	4.05 (1.05)	0.35	.71
Provided information was satisfactory	4.15 (0.08)	3.95 (1.08)	4.25 (0.67)	4.20 (0.77)	0.84	.43
Provided information was confusing	1.48 (0.84)	1.84 (1.12)	1.25 (0.57)	1.50 (0.82)	3.31	.05
Provided information was too much	1.63 (1.06)	2.47 (1.22)	1.47 (0.98)	1.10 (0.31)	11.61	.01
I now understand the prostate cancer diagnosis	3.81 (1.10)	3.84 (1.16)	3.90 (0.94)	3.63 (1.30)	0.34	.62
I now understand prostate cancer treatment	3.59 (1.22)	3.84 (1.17)	3.57 (1.23)	3.37 (1.30)	0.70	.50
I now understand prostate cancer side effects	3.89 (1.01)	3.84 (1.07)	3.90 (0.98)	3.90 (1.04)	0.03	.98
I now understand follow-up care	3.78 (0.99)	3.78 (1.11)	3.87 (0.97)	3.63 (0.96)	0.32	.93
I now understand clinical trials	3.23 (1.24)	3.47 (1.42)	3.27 (1.19)	2.80 (1.11)	1.24	.30
Made me consider more questions to ask	3.73 (0.91)	3.74 (0.81)	3.72 (0.89)	3.75 (1.07)	0.01	.99
Made me seek more information about prostate cancer	3.45 (1.0)	3.63 (0.93)	3.41 (1.07)	3.35 (0.90)	0.45	.95
Enhancing decision making , mean (SD)						2
The information is helpful in decision making	3.56 (1.38)	1.79 (0.92)	4.29 (0.64)	4.10 (1.07)	55.62	.01
Made me think about my treatment choices	3.90 (0.88)	3.89 (0.81)	3.88 (0.7)	3.95 (1.10)	0.04	.96
Made it difficult for me to decide	2.46 (0.91)	2.68 (1.11)	2.39 (0.92)	2.35 (0.93)	0.71	.94
Calmed my nerves about my decision	3.10 (1.94)	2.68 (1.06)	3.12 (0.83)	3.46 (0.89)	3.46	.04

Usability assessment items	Total sample (n = 72)	Study group comparisons				
		Control condition (n = 19, 26%)	PIES[a] with tailoring (n = 32, 45%)	PIES without tailoring (n = 21, 29%)	F_2	*P* value
Made me more anxious about my decision	2.66 (1.18)	3.62 (1.05)	2.45 (1.09)	2.40 (1.27)	3.74	.03
Made the treatment options clear for me	3.58 (0.68)	3.28 (1.27)	3.72 (0.73)	3.63 (0.76)	1.75	.22

[a] Prostate Interactive Educational System.

Differences Between the Tailored and Nontailored PIES Program Evaluation

We found no differences between the tailored and nontailored PIES group among the following group of variables: demographic and clinical factors, PIES evaluation and decision variables, and monitoring style (see Table 3). Thus, data of these 2 groups were combined in subsequent analyses and compared with data from the comparison group (ie, control condition).

Table 3. Prostate Interactive Educational System (PIES)-specific evaluation of the library materials at postintervention assessment (t2) among tailored and nontailored PIES groups[a]

Usability assessment items	Study group comparisons			
	PIES with tailoring (n = 32)	PIES without tailoring (n = 21)	t_{42}	*P* value
PIES information in library is clearly presented	3.77 (0.97)	3.75 (0.80)	0.29	.77
PIES includes everything I need to know	3.67 (1.09)	3.25 (0.91)	1.18	.25
Information is more than I want to know	2.53 (1.01)	2.60 (0.88)	0.35	.73
Graphics are clear	4.03 (0.88)	4.05 (0.69)	0.31	.76
Glossary is helpful	3.93 (0.93)	3.80 (0.73)	0.41	.68
Library was easy to understand	3.96 (1.00)	3.63 (0.82)	1.50	.14
Library provided all the information I need	3.56 (1.09)	3.32 (1.27)	0.60	.55
Library helped me with the decision	3.48 (0.96)	4.47 (0.77)	0.13	.09
Library has more information than what I want	2.68 (1.06)	2.58 (0.96)	0.22	.82

[a] Data are mean (SD).

Examining Preliminary Outcomes: Decisional Variables

We compared decisional process variables between the combined PIES groups and the control group immediately following the intervention (ie, at t2). Results indicated that patients in the PIES groups were significantly more confident about their treatment preferences and significantly less likely to report that they needed more information to make a decision. Moreover, patients in the PIES group indicated that they would need less time to deliberate their treatment options and felt better informed about their choices and potential side effects (see Table 4). Despite these significant differences in decisional process variables, treatment preferences remained stable between the PIES and control group, and no significant impact of PIES on treatment preferences was found (see Table 4).

Table 4. Decisional outcomes at postintervention assessment (t2), by study group

Treatment preferences/decision assessment items		Total sample (n = 72)	Study group comparisons				
			Control condition (n = 19, 26%)	PIES[a] with/without tailoring (n = 53, 74%)	t/χ^2	Degrees of freedom	*P* value
Decisional variables at t2, mean (SD)							
	Have confidence in treatment decision made	3.69 (1.10)	3.22 (1.32)	3.85 (1.022)	–2.35	68	.02
	Preferred more time to think about options	2.52 (1.41)	3.00 (1.41)	2.33 (1.42)	1.81	68	.07
	Preferred more information about prostate cancer	2.77 (1.5)	3.44 (1.54)	2.52 (1.49)	2.48	68	.02
	Feeling informed about prostate cancer and treatment	3.57 (0.91)	3.28 (1.07)	3.74 (0.089)	–1.65	68	.10
Treatment preference at t2 , n/N (%)							
	Surgery	14/65 (22%)	0/16 (0%)	14/49 (29%)	7.3	4	.11
	External beam radiation therapy	11/65 (17%)	5/16 (31%)	6/49 (12%)			
	Brachytherapy	28/65 (43%)	8/16 (50%)	20/49 (41%)			
	Watchful waiting/active surveillance	4/65 (6%)	1/16 (6%)	3/49 (6%)			
	Other	8/65 (12%)	2/16 (13%)	6/49 (12%)			

[a] Prostate Interactive Educational System.

Discussion

Making a treatment decision under conditions of heightened uncertainty, due to the absence of an evidence-based criterion standard treatment option, is difficult and may lead to increased distress, difficulty making a treatment choice, and feelings of decisional regret, especially when treatment outcomes decrease patients' quality of life. Involving prostate cancer patients in treatment decision making has been repeatedly advocated [25]; however, evidence has shown that patients often have difficulties processing treatment-related information, especially in an emotionally charged situation following a cancer diagnosis. We designed the PIES program to address this issue. Our findings showed that patients in the PIES groups, compared with those in the control group were significantly more likely to report higher levels of confidence in their treatment choices, to rate the helpfulness of the information significantly higher for making a treatment decision, and to indicate that the information was emotionally reassuring. Additionally, the participants in the PIES groups thought the information was clear and understandable, were significantly less likely to report a need for more information, and were less anxious about their treatment choices.

Enrollment and Attrition Analyses

Following previous research assessing the feasibility of an applied intervention [18,25], we examined (1) *recruitment* and *retention*, (2) the *acceptability* of PIES, and (3) *preliminary outcomes*. The high acceptance and completion rate of the baseline questionnaire (80%) suggested that patients participating in the study had a chance to explore the PIES program and found it acceptable.

Evaluation of PIES and Print Materials

Our study design included a time and attention comparison condition, the methodologically most rigorous approach to test a novel intervention. Patients in the control condition were asked to read 2 NCI-published brochures that provide extensive, albeit noninteractive, information about prostate cancer and its treatment. The comprehensive nature of the provided information might be responsible for the overall increase in understanding among patients in the control group. Despite a uniform increase in being informed, patients completing PIES were significantly less confused about their treatment options,

and felt that the information they received was the "right" amount and was emotionally reassuring. The latter point is particularly noteworthy, as it underscores the importance of addressing the inherently distressing nature of a prostate cancer diagnosis. Guided by our theoretical self-regulation framework that incorporates both cognitive and emotional processing of health-relevant information, we addressed the emotional aspect of prostate cancer throughout the PIES program. Survivors talked freely about the emotional impact of prostate cancer on themselves and their family, and how they coped with the diagnosis. Physicians attempted to emotionally reassure patients by providing normalizing statements (eg, "you are not alone in this diagnosis, thousands of men are being diagnosed with prostate cancer"). They also made references to available support from psychologists and social workers, and discussed medical solutions to erectile and urinary dysfunction. Although the program needs to be evaluated further to relate specific components of PIES usage to specific decision and adjustment variables, the overall results support our comprehensive cognitive–affective approach to patient information.

Differences Between the Tailored and Nontailored PIES Program Evaluation

The lack of an effect of the tailoring variable was surprising. There are two potential reasons for this outcome. First, we only tailored written information contained in the library books; all physician answers and patient stories were nontailored. As patients were free to explore the program at will, some might have spent more time with patient stories and physician answers, rather than reading the tailored written information contained in the library. Second, patients might have self-tailored their information intake by acquiring the amount and detail of information that corresponded to their monitoring style. As exploration of the software was unstructured, it is reasonable to assume that they would explore those issues that were relevant and interesting to them and avoid those items that might be uninteresting or anxiety provoking. It is therefore possible that patients' self-tailoring exploration behavior superseded the tailoring capabilities of the software, particularly if patients did not spend enough time in the library reading the tailored materials. Although preliminary results from our study suggest that patients might engage in self-tailoring activities, the issue of tailoring to information processing within a multimedia environment needs to be examined in more detail before a comprehensive recommendation about tailoring to this variable can be made. In addition to demonstrating acceptability on the patient side, we further demonstrated that implementation into a clinical consultation service is possible. Despite some time constraints that did not allow all patients to complete the postintervention questionnaire, a large majority completed all assessments. This problem could be mitigated in future implementations by making the program available to patients via the Internet prior to their physician appointment. It would be simple to provide patients with the program's Internet link and an individualized login and password to access the software, when they make an appointment to see the physician. The next step then would be to link the program with the electronic medical record system that could transmit patients' concerns and preferences to the treating physician prior to the consultation. This would give physicians valuable information about the upcoming consultation, and would allow them to tailor the information provided more closely to patients' needs.

Preliminary Outcomes: Decisional Variables

The results of the study indicate that PIES improved some of the decisional process variables under study. Specifically, PIES succeeded in satisfying the patients' information needs, increased their confidence about their treatment preferences, and provided the information in an emotionally reassuring way. PIES also increased patients' knowledge about their treatment options and associated side effects. These results are particularly promising given that they were obtained within the framework of a time and attention comparison condition (ie, control group), which provided considerably more information than standard or usual care. Patients receiving standard or usual care might or might not receive written information to take home in addition to the physician consultation. In the present study patients received 2 brochures and were asked to spend 45 minutes reading both of them; thus, they were most likely better informed than the average patient attending a physician consultation. We would expect, therefore, that PIES would perform even better when compared with standard or usual care.

As expected, PIES did not change patients' treatment choice. Because PIES was designed with the intention to inform patients and to help them identify what is important to them with regard to treatment outcomes and future quality of life, the lack of a significant effect on change in treatment choices was not unexpected and indeed affirmed our approach.

No study is without limitations and the present one is no exception. The study findings confirmed PIES' acceptability and demonstrated its ability to influence important decisional variables. Although the results are not definitive, documenting acceptability and feasibility and preliminary results of its effect is an important first step before proceeding to a larger randomized controlled trial. Second, our patient population was not representative of the general population, as the majority were well educated (56% college educated or higher). Third, we examined study outcomes before and after the completion of PIES intervention. Examining the impact of the intervention several weeks later might reveal a stronger impact of the intervention on psychosocial outcomes such as regret, recurrence worries, and psychological stress.

In sum, the present study provides evidence that PIES is acceptable to patients, can be implemented into a routine clinic program, successfully improves patients' knowledge about treatments and side effects, and increases their confidence in treatment decision making.

Acknowledgments

Supported by grants from the National Cancer Institute (CA6136-05, CA90904-02) and the Department of Defense (DAMD W81 XWH-04-1-0179).

We wish to thank Jessica Lake for her extensive assistance with manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The virtual health center.

[PNG File, 1MB - jmir_v14i1e6_app1.png]

Multimedia Appendix 2

The Library.

[PNG File, 1MB - jmir_v14i1e6_app2.png]

Multimedia Appendix 3

The physicians' offices.

[PNG File, 1MB - jmir_v14i1e6_app3.png]

Multimedia Appendix 4

The support group room.

[PNG File, 748KB - jmir_v14i1e6_app4.png]

References

1. National Cancer Institute. 2011. Prostate Cancer URL: http://www.cancer.gov/cancertopics/types/prostate [accessed 2011-06-20] [WebCite Cache ID 5zaNWGbAy]
2. Logothetis CJ, Millikan R. Update: NCCN practice guidelines for the treatment of prostate cancer. Oncology (Williston Park) 1999;13(11A):118-132.
3. Bahnson RR, Hanks GE, Huben RP, Kantoff P, Kozlowski JM, Kuettel M, National Comprehensive Cancer Network. NCCN Practice Guidelines for Prostate Cancer. Oncology (Williston Park) 2000 Nov;14(11A):111-119. [Medline: 11195405]
4. Talcott JA, Manola J, Clark JA, Kaplan I, Beard CJ, Mitchell SP, et al. Time course and predictors of symptoms after primary prostate cancer therapy. J Clin Oncol 2003 Nov 1;21(21):3979-3986. [doi: 10.1200/JCO.2003.01.199] [Medline: 14581420]
5. Litwin MS, Gore JL, Kwan L, Brandeis JM, Lee SP, Withers HR, et al. Quality of life after surgery, external beam irradiation, or brachytherapy for early-stage prostate cancer. Cancer 2007 Jun 1;109(11):2239-2247 [FREE Full text] [doi: 10.1002/cncr.22676] [Medline: 17455209]
6. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. JAMA 2005 May 4;293(17):2095-2101 [FREE Full text] [doi: 10.1001/jama.293.17.2095] [Medline: 15870412]
7. Diefenbach MA, Mohamed NE. Regret of treatment decision and its association with disease-specific quality of life following prostate cancer treatment. Cancer Invest 2007 Sep;25(6):449-457. [doi: 10.1080/07357900701359460] [Medline: 17882657]
8. O'Connor AM, Bennett CL, Stacey D, Barry M, Col NF, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2009(3):CD001431. [doi: 10.1002/14651858.CD001431.pub2] [Medline: 19588325]
9. DePalma A. Prostate Cancer Shared Decision: a CD-ROM educational and decision-assisting tool for men with prostate cancer. Semin Urol Oncol 2000 Aug;18(3):178-181. [Medline: 10975488]
10. Kim SP, Knight SJ, Tomori C, Colella KM, Schoor RA, Shih L, et al. Health literacy and shared decision making for prostate cancer patients with low socioeconomic status. Cancer Invest 2001;19(7):684-691. [Medline: 11577809]
11. Diefenbach MA, Butz BP. A multimedia interactive education system for prostate cancer patients: development and preliminary evaluation. J Med Internet Res 2004 Jan 21;6(1):e3 [FREE Full text] [doi: 10.2196/jmir.6.1.e3] [Medline: 15111269]
12. Leventhal H. Findings and theory in the study of fear communications. Adv Exp Soc Psychol 1970;5:119-186. [doi: 10.1016/S0065-2601(08)60091-X]

13. Leventhal H, Diefenbach M, Leventhal EA. Illness cognition: using common sense to understand treatment adherence and affect cognition interactions. Cognitive Therapy and Research 1992;16(2):143-163. [doi: 10.1007/BF01173486]
14. Peters E, Diefenbach MA, Hess TM, Västfjäll D. Age differences in dual information-processing modes: implications for cancer decision making. Cancer 2008 Dec 15;113(12 Suppl):3556-3567 [FREE Full text] [doi: 10.1002/cncr.23944] [Medline: 19058148]
15. Miller SM. Monitoring versus blunting styles of coping with cancer influence the information patients want and need about their disease. Implications for cancer screening and management. Cancer 1995 Jul 15;76(2):167-177. [Medline: 8625088]
16. Skinner CS, Strecher VJ, Hospers H. Physicians' recommendations for mammography: do tailored messages make a difference? Am J Public Health 1994 Jan;84(1):43-49. [Medline: 8279610]
17. Miller DL, Manne SL, Taylor K, Keates J, Dougherty J. Psychological distress and well-being in advanced cancer: the effects of optimism and coping. J Clin Psychol Med Settings 1996;3(2):115-130. [doi: 10.1007/BF01996132]
18. Kazak AE, Simms S, Alderfer MA, Rourke MT, Crump T, McClure K, et al. Feasibility and preliminary outcomes from a pilot study of a brief psychological intervention for families of children newly diagnosed with cancer. J Pediatr Psychol 2005 Dec;30(8):644-655. [doi: 10.1093/jpepsy/jsi051] [Medline: 16260434]
19. Steptoe A. An abbreviated version of the Miller Behavioral Style Scale. Br J Clin Psychol 1989 May;28 (Pt 2):183-184. [Medline: 2743058]
20. NIH. Us Too International, Inc, National Cancer Institute. Bethesda, MD: NIH Publication; 2000. Understanding Treatment Choices for Prostate Cancer (NIH publication no. 00-4659) URL: http://www.cancer.gov/images/Documents/5dc60312-9a4d-4c08-8d0e-df0731287723/choices.pdf [accessed 2011-12-23] [WebCite Cache ID 649IZJIKp]
21. National Cancer Institute. Bethesda: NIH; 2008 Nov 20. What You Need to Know About Prostate Cancer (NIH publication no. 08-1576) URL: http://www.cancer.gov/cancertopics/wyntk/prostate/page1/AllPages [accessed 2011-12-23] [WebCite Cache ID 649IOv5em]
22. NIH. National Cancer Institute. Bethesda, MD: NIH Publication; 1989 Apr. Making Health Communication Programs Work (NIH publication no. 89-1493) URL: http://www.cancer.gov/cancertopics/cancerlibrary/pinkbook/Pink_Book.pdf [accessed 2011-12-23] [WebCite Cache ID 649IfMG9P]
23. Robinson TN, Patrick K, Eng TR, Gustafson D. An evidence-based approach to interactive health communication: a challenge to medicine in the information age. Science Panel on Interactive Communication and Health. JAMA 1998 Oct 14;280(14):1264-1269. [Medline: 9786378]
24. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. Psychosom Med 1979 May;41(3):209-218 [FREE Full text] [Medline: 472086]
25. Jibaja-Weiss ML, Volk RJ, Granch TS, Nefe NE, Spann SJ, Aoki N, et al. Entertainment education for informed breast cancer treatment decisions in low-literate women: development and initial evaluation of a patient decision aid. J Cancer Educ 2006;21(3):133-139. [doi: 10.1207/s15430154jce2103_8] [Medline: 17371175]

Abbreviations

IES-R: Impact of Event Scale-Revised
NCI: National Cancer Institute
PIES: Prostate Interactive Educational System
PSA: prostate-specific antigen

Edited by G Eysenbach; submitted 18.07.11; peer-reviewed by M Morris, D Bowen; comments to author 26.08.11; revised version received 26.09.11; accepted 27.09.11; published 13.01.12.

Please cite as:
Diefenbach MA, Mohamed NE, Butz BP, Bar-Chama N, Stock R, Cesaretti J, Hassan W, Samadi D, Hall SJ
Acceptability and Preliminary Feasibility of an Internet/CD-ROM-Based Education and Decision Program for Early-Stage Prostate Cancer Patients: Randomized Pilot Study
J Med Internet Res 2012;14(1):e6
URL: http://www.jmir.org/2012/1/e6/
doi:10.2196/jmir.1891
PMID:22246148

Original Paper

Results of an Online Community Needs Assessment for Psychoeducational Interventions Among Partners of Hereditary Breast Cancer Previvors and Survivors

Kenneth P Tercyak[1], PhD; Darren Mays[1], PhD, MPH; Tiffani A DeMarco[1], MS; McKane E Sharff[1], BA; Susan Friedman[2], DVM

[1]Georgetown University Medical Center, Washington, DC, United States
[2]FORCE, Inc., Tampa, FL, United States

Corresponding Author:
Kenneth P Tercyak, PhD
Georgetown University Medical Center
Suite 4100
3300 Whitehaven Street, NW
Washington, DC,
United States
Phone: 1 202 687 0802
Fax: 1 202 687 8444
Email: tercyakk@georgetown.edu

Abstract

Background: Spouses and partners ("partners") of women at-risk for ("previvors") and surviving with hereditary breast/ovarian cancer are a primary source of support within their families. Yet, little is known about partners' needs for psychoeducational intervention to enhance their cancer risk knowledge, coping, and support role functioning.

Objective: To determine the type and range of need for psychoeducational intervention among partners of hereditary breast cancer previving and surviving women, and to understand the potential role of the Internet and other communication channels in meeting that need.

Methods: We conducted a secondary data analysis on partners' needs that were originally assessed via an online community-based organization devoted to hereditary breast cancer. Partners' demographic characteristics, need for psychoeducation, and likelihood of using various communication channels were assessed along with other constructs. Analyses examined commonly-occurring clusters of likely intervention use and by communication channel.

Results: Partners (n =143) endorsed a moderately high level of need for psychoeducation and did so across multiple content areas (e.g., role functioning, decision making, communication, intimacy). Factor analysis identified three commonly-preferred communication channels: 1) self-help materials, 2) online interactions, and 3) interpersonal interactions. A cluster analysis among these factors identified three groups of partners based on their likelihood of psychoeducational intervention use (low [18%], moderate [55%], and high [27%] users). In a covariate-adjusted MANOVA, moderate and high intervention users reported significantly greater need for psychoeducation compared to low users ($F_{2,132} = 9.15$, $P < .001$).

Conclusions: A majority of assessed partners perceived a need for psychoeducational interventions surrounding hereditary breast cancer risk. Internet-based, interactive resources may be an efficient mechanism to reach large numbers of partners with tailored content. Research is warranted to inform the design and deployment of these resources to ensure quality and high impact, and ultimately to examine ways to integrate these resources into clinical care.

(J Med Internet Res 2012;14(1):e15) doi:10.2196/jmir.1847

KEYWORDS

Breast cancer; hereditary cancer; social support; psychoeducation; psychosocial intervention

Introduction

Every year in the United States, breast cancer is diagnosed in over 200,000 women [1]: internationally, it is the second most commonly diagnosed malignancy and the leading cause of cancer-related death among women [2]. Of these cases, approximately 5%–10% are hereditary [3]. Most cases of hereditary breast cancer are attributable to germline mutations in one of two major breast cancer-predisposing genes, *BRCA1* or *BRCA2* (*BRCA1/2*) [4,5]. Women with a *BRCA1/2* mutation face up to an 85% lifetime risk for breast cancer and up to a 65% lifetime risk for ovarian cancer [4,5]. Moreover, these cancers are often diagnosed in women at younger ages than average [6]. Importantly, when a *BRCA1/2* mutation is identified in an individual, there is a 50% chance that first-degree relatives (eg, male and female children and siblings) have also inherited the mutation and may therefore face increased risks for cancer [5]. For women with a known *BRCA1/2* mutation, breast cancer screening consists of mammography and breast magnetic resonance imaging starting at age 25 years [7,8]. Breast cancer risk-reduction options include chemoprevention, prophylactic mastectomy, and prophylactic oophorectomy, or a combination of these [7,8]. Prophylactic oophorectomy is recommended after childbearing is completed to reduce mortality from ovarian cancer. Given all of these considerations, the presence of hereditary cancer confronts families with many complex, emotionally charged decisions, and increased awareness of familial cancer susceptibility brings about a lifelong impact [9-11].

Genetic counseling and testing for *BRCA1/2* mutations is a well-established component of the identification and management of hereditary breast and ovarian cancer syndrome among those who are at risk [12,13]. Though cancer care providers (eg, genetic counselors, nurses, oncologists, and surgeons) are a common source of medical support for those who undergo genetic testing, women's family members, especially their partners, are the most likely source of psychosocial support [14,15]. Indeed, prior work has demonstrated that women's *BRCA1/2* test-related decisions are often discussed with their partners, and most women feel supported by their partners [16,17]. However, these same data also indicate that, in the face of less support and greater protective buffering in partnered couples (ie, hiding worries, denying concerns, and engaging in avoidant behaviors), poorer psychological outcomes can ensue [10,18,19]. By contrast, greater partner support predicts better psychological outcomes among these dyads [16].

In light of this, it is critical that families facing the risk of hereditary breast and ovarian cancer be adequately supported and empowered, both medically and psychosocially, before and after learning about their disease risk [19-21]. Given the limited time and resources of most cancer care providers to offer ongoing psychosocial and educational support to women tested for *BRCA1/2* mutations and their family members, it is imperative that adjunctive models of psychoeducational support be offered outside of the health care setting to better meet the needs of women who are at risk of familial breast cancer and their partners. Psychoeducation, which is a well-known intervention model for providing informational and psychosocial support for chronically ill women and their partners [22-25], may be an important intervention method for families facing the risk of hereditary breast and ovarian cancer as an adjunct to standard cancer care and cancer prevention.

The Internet is a primary resource for those seeking information and support about cancer [26,27]. Internet-based resources are particularly valuable tools for those facing a risk of cancer and other chronic diseases, as they can provide timely, relevant resources [28-30]. Internet-based resources are also commonly available for persons with a known risk of hereditary breast cancer. For example, the National Cancer Institute, the American Cancer Society, Susan G. Komen for the Cure, and other leading breast cancer advocacy groups sponsor websites devoted to educating the public about hereditary breast cancer, prevention, treatment, and related issues. However, and despite the familial nature of hereditary breast cancer and the involvement of relatives in genetic counseling and testing [16], Internet-based psychoeducation has not been developed specifically for partners of women surviving with hereditary breast cancer and those at risk but who have not developed disease (ie, previvors).

Partners of previvors and survivors, especially male partners, may prefer Internet-based resources to face-to-face psychoeducation because they offer anonymity surrounding sensitive topics and emotional experiences, and provide direct access to needed information [30-33]. It is likely that partners (and male partners in particular) have specific and unique needs for psychoeducation that could assist them in supporting previvors and survivors, which may include education about hereditary breast cancer, helping facilitate decisions about *BRCA1/2* genetic counseling and testing and options after testing, establishing and maintaining open communication within the partnership about hereditary breast cancer and related concerns, and performing supportive behaviors and managing stress and uncertainty in the face of previvors' and survivors' hereditary cancer risk and risk of cancer in the family [32,34]. To date, however, there has been no systematic examination of the psychosocial support needs of partners of previvors and survivors of hereditary breast cancer.

To address this gap, we report on the results of an online community needs assessment conducted with this target population. Specifically, the assessment, which was conducted by Facing Our Risk of Cancer Empowered (FORCE), sought to describe the need for psychoeducational interventions that could be offered via the Internet and other communication channels among partners of hereditary breast cancer previvors and survivors. FORCE maintains in-person, telephone, and Web-based programs that provide information, peer support, resources, and a community tailored to individuals at high genetic risk. Its website (www.facingourrisk.org) is the leading site specifically devoted to the community of hereditary breast cancer previvors and survivors. It is expected that the findings of this assessment would be used to inform the planning and development of new interactive and Internet-based psychoeducational interventions for this target population.

Methods

Overview

This is a secondary analysis of data originally collected by FORCE through a Web-based survey to determine its online community members' needs. The work was guided by the PRECEDE portion of the PRECEDE-PROCEED conceptual framework for designing health promotion programs [35,36]. Briefly, the PRECEDE framework refers to Predisposing, Reinforcing, and Enabling factors in educational diagnosis and evaluation [36]. According to PRECEDE, a critical initial step in planning health promotion and intervention programs is to understand gaps between resources that are currently available within the target community and community members' perceived needs for additional resources [35,36].

Setting

The Internet survey was conducted by FORCE, Inc. Based in Tampa, FL, USA, FORCE is a national 501(C)3 not-for-profit organization devoted to raising awareness about hereditary breast and ovarian cancer and *BRCA1/2* mutations. The FORCE website is the foremost lay Internet site devoted to the cancer education and support needs of persons with or at risk for hereditary breast cancer and their family members. The FORCE website contains timely and accurate information about hereditary cancer, cancer risk assessment, and other related topics. In addition to educational information, resources offered by the organization include national and local outreach groups, online webinars, print brochures, a toll-free helpline, periodic newsletters, and an annual educational meeting.

Recruitment and Data Collection

The needs assessment sample consisted of 143 partners who responded to the Web-based survey. The anonymous survey took approximately 10 minutes to complete, and no personally identifying information was collected. The survey was made available via the FORCE website homepage from November 2010 to February 2011. The heading for the survey targeted spouses or partners of women who were at risk for hereditary breast cancer, defined specifically as having *BRCA1/2* genetic mutations or a family history of breast cancer, or who had a *BRCA1/2*-linked cancer. Respondents were asked to affirm the following statement prior to completing the survey:

> *Participation in this survey is limited to spouses and partners (men and women, age 18 or older) of women with a BRCA mutation or family history of cancer. By continuing, you are agreeing with the above terms and volunteering to participate. If you do not wish to participate, please close your browser or exit at any time.*

The protocol for this secondary data analysis was reviewed and approved by the Institutional Review Board at Georgetown University.

Measures

Demographics

The demographic characteristics assessed were respondent age, gender, race/ethnicity, and highest level of education attained. In addition, 2 items assessed whether respondents had any children, and whether they had any female children.

Clinical Characteristics

Clinical characteristics assessed included whether respondents' partners had a diagnosis of breast cancer, had been tested for a *BRCA1/2* genetic mutation, and had surgery to remove her breasts (ie, prophylactic mastectomy) or ovaries (ie, prophylactic oophorectomy), or for breast reconstruction. Based on these items, 3 variables were created to indicate whether each respondent's partner had (1) a diagnosis of breast cancer, (2) received *BRCA1/2* genetic testing, and (3) undergone any of the 3 surgery types we inquired about.

Internet and Email Usage

Because FORCE has a large and active online community, respondents' use of technology was presumed to be moderately high. However, to evaluate this presumption among partners, Internet use was formally assessed using 2 items. The first item asked "How often do you go online to access the Internet?" with response options ranging on a 4-point Likert-type scale from never to very often (more than once/week). The second item asked "When you go online, where do you primarily access the Internet from?" with response options including home, work, Internet café, family members or friends' home, and other. Email use was assessed by asking "Do you have an email address for your personal use?" Respondents were dichotomized as high Internet users if they accessed the Internet very often, did so from home, and had an email address for personal use; all other participants were categorized as low Internet users [37].

Preferred Psychoeducational Content

We used 7 items to assess respondents' preferences for psychoeducational content. From the perspective of partners, topics queried included understanding my role/knowing what to expect, communicating with my spouse/partner, helping my spouse/partner make decisions, communicating with adult relatives, communicating with children, intimacy after diagnosis or surgery, and speaking with other spouses/partners going through a similar situation. Response options for each item were yes, no, and I don't know. Principal components factor analysis of these 7 items confirmed a single-factor solution (eigenvalue = 2.93). Items were analyzed individually and an overall content score was also created by summing responses to the 3 items where yes, I don't know, and no were again assigned a value of 2, 1, or 0, respectively (range 0–14, Cronbach alpha = .75). As such, higher scores reflected stronger preferences for more psychoeducational content.

Communication Channels

We used 11 items to assess respondents' likelihood of using psychoeducational resources offered through the following communication channels: regular mail, toll-free telephone line, email, national and local FORCE meetings, printed booklet or guide, periodic newsletter, expert teleconference or webinar,

video or DVD, Web-based message board, and Web-based chat. Items were preceded by a statement instructing respondents to indicate how likely they would be the find the following resources and information for spouses or partners to be useful. Response options were based on a 7-point Likert-type scale with anchors at values of 1 (not at all likely) and 7 (very likely).

These 11 items were subsequently factor-analyzed to empirically derive subscales with eigenvalues >1: we identified 3 subscales. Subscale scores were then created by averaging responses to the items loading on each subscale to ensure that all subscales were based on a common underlying metric. The *self-help* subscale consisted of 4 items assessing the likelihood of using a printed booklet or guide, newsletter, and materials delivered via postal mail and email (eigenvalue = 4.8, Cronbach alpha = .82, mean 5.1, SD 1.4, range 1–7). The *online interaction* subscale consisted of 4 items assessing the likelihood of using resources offered via a Web-based message board or chat, embedded video or DVD, and expert webinar or teleconference (eigenvalue = 1.6, Cronbach alpha = .82, mean 4.2, SD 1.5, range 1–7). Finally, the *interpersonal interaction* subscale consisted of 3 items assessing the likelihood of using in-person resources offered via national and local FORCE meetings, as well as a telephone hotline (eigenvalue = 1.1, Cronbach alpha = .76, mean 3.6, SD 1.6, range 1–7).

Perceived Need for Psychoeducation

We used 3 items to assess respondents' perceived need for psychoeducation. The items were introduced to respondents with a brief description of the informational resources that could be made available through FORCE, followed by items assessing whether respondents (1) perceived a *need* for more resources or support, (2) *wanted* more resources or support, and (3) would *use* more resources or support. Response options for each item were yes, no, and I don't know. An overall score was created by summing responses to the 3 items with assigned values of yes (2), I don't know (1), and no (0) (range 0–6, Cronbach alpha = .70): higher scores reflect greater need.

Data Analysis

Statistical analyses were conducted in several steps. First, we used descriptive statistics to characterize the study sample and describe their preferences and perceived needs for psychoeducation. Second, we subjected communication channel subscale scores (ie, self-help, online interaction, and interpersonal interaction) to a hierarchical cluster analysis to determine first-order groupings of respondents based on their self-reported likelihood of using resources offered through specific communication channels. We selected the unweighted pair group method using arithmetic averages, which defines clusters based on the average pairwise proximities between clusters of all pairs of observations [38,39]. This method has performed well under various conditions in Monte Carlo simulation studies and was suitable for the data [39]. The analysis produced 3 clusters of potential users of psychoeducational interventions with eigenvalues >1. To confirm the validity of these groupings, we conducted pairwise comparisons of mean subscale scores for preferred communication channels across potential user clusters, applying the Tukey post hoc adjustment for multiple comparisons [40].

Subsequently, bivariate tests (ie, F tests, t tests, χ^2 tests) examined relationships between demographic and clinical characteristics, and perceived need for psychoeducation, across user clusters. Finally, a multivariate analysis of variance (MANOVA) examined variability in need across users, adjusting for participant demographic characteristics (ie, age, gender, race/ethnicity, any children) as covariates [40].

Prior to analyses, we studied patterns of missing data in focal variables, including preferred communication channels, psychoeducational content, and need for support. While data were missing for any one of these variables for only a few participants (ie, ≤10%), Little's [41] χ^2 test for data missing completely at random indicated the presence of identifiable patterns of missing data ($\chi^2_{32} = 56.7$, $P = .005$). To account for missing data on these variables, we used a single regression-based imputation method, imputing predicted values based on demographic characteristics including age, gender, and race/ethnicity. This method has been shown to be adequate for imputing missing values when the proportion of participants with missing data is relatively low, as was the case for our sample [40].

Results

Study Sample

Characteristics of the respondent sample (n = 143) are displayed in Table 1. Respondents averaged 45.8 years of age. A majority were male (86.0%), white (94.4%), had a college education or higher (86.7%), and had 1 or more children (69.2%). Most respondents reported that their spouse/partner had *BRCA1/2* genetic testing previously (91.6%) and some form of surgery (78.3%). Most respondents (55%) were categorized as high-level Internet users, and their need for psychoeducation was moderately high (mean 4.6, range 0–6). Descriptive information for partners' preferred psychoeducational content is displayed in Table 2. As shown, there was uniformly strong interest in all content areas presented, and particularly for content focusing on the partner's role and knowing what to expect, decision making, communication, and intimacy.

Table 1. Sample characteristics (n = 143)

Demographics	Mean	SD	n	%
Age (years)	45.8	10.5		
Gender				
Male			123	86.0
Female			20	14
Race/ethnicity				
White			135	94.4
Nonwhite			8	6
Education				
< College			19	13
≥ College			124	86.7
Family characteristics				
≥1 Child			99	69
0 Children			43	30
≥1 Female child			73	51
0 Female children			70	49
Clinical characteristics				
Breast cancer diagnosis			48	34
No			91	64
BRCA testing			131	91.6
No			12	8
Breast/ovarian surgery			112	78.3
No			31	22
Internet use				
High			78	55
Low			65	45
Need for psychoeducation (range 0–6, alpha = .77)	4.6	1.8		
Communication channel [a]				
Self-help (range 0–7, alpha = .82)	5.1	1.4		
Online interaction (range 0–7, alpha = .82)	4.2	1.5		
Interpersonal (range 0–7, alpha = .76)	3.6	1.6		

[a] Scores based on the average response to items within each subscale based on a 7-point Likert-type scale with anchors at values for 1 (not at all likely) and 7 (very likely).

Table 2. Preferred psychoeducational content

Topic [a]	Mean	SD
Understanding my role/knowing what to expect	1.89	0.45
Helping my partner make decisions	1.75	0.63
Communicating with my spouse/partner	1.68	0.73
Intimacy after diagnosis/surgery	1.67	0.73
Speaking with others undergoing a similar experience	1.47	0.82
Communicating with children	1.45	0.84
Communicating with adult relatives	1.24	0.90

[a] Response options for each item were yes, I don't know, and no and assigned values of 2, 1, or 0, respectively.

Cluster Analysis

The cluster analysis of communication channels identified 3 distinct groups of partners based on their likelihood of using psychoeducational interventions (Table 3). The smallest proportion of participants (18%, n = 26) fell into the low-use cluster (eigenvalue = 1.36), which was characterized by the lowest likelihood of their expected use of all 3 resource types. A majority of respondents (55%, n = 78) were characterized by the moderate-use cluster (eigenvalue = 4.58), which included intermediate levels of need for each resource type. Finally, just over one-quarter (27%, n = 39) of partners were classified in the high-use group (eigenvalue 1.04), with the highest average likelihood of using all 3 communication channels. All pairwise mean comparisons of the likelihood of using self-help, online, and interpersonal interaction resources were significantly different across psychoeducational intervention use clusters at $P < .05$ using the Tukey post hoc adjustment, confirming cluster validity [38].

Table 3. Cluster analysis of partners' preferred communication channels

	Likelihood of psychoeducational intervention use (clusters)					
	Low (26/143, 18%)		Moderate (78/143, 55%)		High (39/143, 27%)	
Communication channel[a]	Mean	SD	Mean	SD	Mean	SD
Self-help	3.5	1.5	5.2	1.1	5.8	1.2
Online interaction	2.1	1.1	4.4	1.0	5.3	1.1
Interpersonal interaction	1.8	0.95	3.1	0.95	5.7	0.70
Eigenvalue	1.36		4.57		1.04	

[a] Values for communication channel are based on a 7-point Likert-type scale with anchors at values for 1 (not at all likely) and 7 (very likely). All pairwise mean comparisons of preferred communication channels across clusters are statistically significant at $P < .001$ using the Tukey post hoc test, except for high and moderate clusters for the self-help channel, where $P = .02$.

Bivariate Relationships

Participants in the low-, moderate-, and high-use clusters did not significantly differ based on demographics, family characteristics, clinical characteristics, or Internet use. Significant differences in partners' perceived need for psychoeducation were evident across clusters. Specifically, those in the moderate-use and high-use clusters reported significantly greater need than did participants in the low-use cluster ($F_{2,142} = 13.3$, $P < .001$).

Multivariate Analysis of Variance

Findings from the MANOVA indicate that, after adjusting for demographic covariates including age, gender, white race, and any children, significant variability existed in partners' need for psychoeducation across clusters (Wilks lambda = 0.88, $F_{2,132} = 9.15$, $P < .001$; group main effect $F_{2,138} = 9.15$, $P < .001$). Examination of pairwise comparisons of adjusted mean need across clusters indicated that partners in the low-use cluster (mean 3.8, SE 0.48) reported significantly less need than partners in the high- (mean 5.7, SE 0.41, $P = .002$) and moderate-use (mean 5.2, SE 0.38, $P < .001$) clusters: need did not differ significantly between partners in the high- and medium-use clusters ($P = .32$).

Discussion

The purpose of this study was to describe and determine the need for psychoeducational interventions among partners of previvors and survivors of hereditary breast cancer, with an emphasis on using the Internet and other remote communication channels to reach a geographically dispersed target population. The findings suggest that partners have a moderately high self-assessed need for psychoeducational interventions, and tend to prefer printed self-help and interactive online resources (though other interpersonal channels received interest as well). With respect to content that would likely resonate with partners, all topics inquired about during the needs assessment received

endorsement and particularly topics normalizing the partner's role as an informed supporter of women facing hereditary cancer risk, with training in coping and communication skills. Interestingly, we were able to empirically derive 3 distinct groups of potential users of psychoeducational interventions based on their communication channel preferences, including those who may be the most and least likely to use self-help, as well as online and interpersonal interaction-based resources. It is expected that these results can inform planning and development of new interactive, Internet-based intervention tools that are specifically directed toward and designed to meet the psychoeducational needs of previvors' and high risk survivors' partners.

Prior studies of the use of Internet-based information and support resources suggest that our findings may, in part, reflect the fact that the needs assessment sample comprised predominantly men. Partners endorsed a high need for psychoeducation, which is consistent with earlier research demonstrating that men are less likely than women to participate in in-person psychosocial cancer support groups, yet express levels of need for resources and support similar to those of women [42,43]. Additionally, the sample of partners indicated a strong need for Internet-based resources, especially partners in the high-need group. Men in particular may be more likely to use Internet-based psychoeducational resources because they allow for anonymous discussion of sensitive topics (eg, intimacy and relationship concerns) [33,44]. Men have also been observed to participate less frequently than women in online discussion forums surrounding cancer, but when they do they tend to express sentiments of fear and anxiety [45] and raise emotional topics, such as the prospect of losing their spouse or partner to cancer [31,32,46]. Moreover, partners of women facing breast cancer experience stress and anxiety when confronting complex issues surrounding breast cancer risk [34,47,48]. Because they themselves may be distressed and in need of information about sensitive topics, partners of breast cancer survivors and previvors may prefer the privacy and anonymity offered by Internet-based psychoeducation for meeting some of their needs.

Our findings also raise important considerations regarding the diversity of partners and the need for psychoeducation targeted toward special populations. While a majority of the sample were men, a significant minority (14%) of partners who responded to the online needs assessment were women. This suggests that when developing resources targeted to partners' needs, one should take into consideration the diversity within this population, including same-sex couples. Evidence regarding the impact of breast cancer on same-sex couples and their need for psychoeducation is scarce [49-51]. Further research is needed to examine potential unique resource needs of same-sex couples as they face the risk of hereditary breast cancer.

Male breast cancer is rare and accounts for only a small proportion of hereditary breast cancer cases (<1%) [52,53]. Men determined to be *BRCA1/2* mutation carriers have elevated risks for breast and prostate cancer, as well as other forms of the disease [54]. Families of male breast cancer previvors and survivors may face similarly complex issues when confronted with decisions surrounding preventive screening and treatment, and the possibility of mutations in offspring and other first-degree relatives [18,55]. Male breast cancer previvors and survivors, and their partners also likely have unique needs for resources to help them navigate these complex issues, which may necessitate targeted interventions and should be investigated in future studies.

Internet-based resources may be ideally suited to address the psychoeducational needs of partners as described herein. Internet-based tools could enable health professionals to effectively incorporate self-help materials that could be downloaded or printed from the Web, as well as leverage interactive multimedia aids such as videos, expert webinars, and online discussion forums to deliver content, provide support, and engage participants [28]. Moreover, Internet-driven intervention-delivery approaches provide ample flexibility to tailor content to individuals' needs and offer targeted materials for population subgroups, which are strategies that have been shown to improve intervention outcomes [56]. Finally, an Internet-based format would provide partners with opportunities for social networking and interaction with other partners, and offer cancer care providers the opportunity to participate in the online space. Such components would attend to the needs of partners preferring to experience more interpersonal connections and communications as well.

Professionally facilitated, Internet-based psychoeducational interventions implemented in diverse populations [30], including those with cancer [28], have shown promising results. Intervention platforms such as the Comprehensive Health Enhancement and Support System improve psychosocial outcomes among women with breast cancer [57-59]. There remains, however, a need to rigorously develop and evaluate Internet-based resources with the specific needs of partners in mind. They are not directly affected by hereditary cancer themselves, yet constitute an important group in the comprehensive cancer care-delivery system. The results of this assessment will be useful to inform efforts to reach out to and support their roles in the future.

Moving forward, it is critical to carefully investigate the ways in which Internet-based psychoeducation and support resources can be integrated into the delivery of health care services for hereditary breast cancer, such as *BRCA1/2* genetic counseling and testing services [30,60]. The Internet is a primary resource for those seeking information about cancer [26,27], and health care providers must acknowledge that patients will use the Internet as a source of information whether it is suggested to them or not. Therefore, it compels providers to take responsibility for directing their patients toward credible, evidence-based resources [61]. Clinical providers should also be aware of important considerations surrounding Internet-based information and support resources, such as the lingering digital divide, which may prevent patients of certain sociodemographic backgrounds (eg, lower socioeconomic status, minority racial/ethnic groups, and the elderly) from having equal access [62,63].

Limitations

The findings reported herein should be interpreted in light of the limitations of this work. The original needs assessment data relied on a convenience sample of partners responding to an

Internet-based survey of unknown denominator (reach). Moreover, our focus was on partners of previving and surviving women, and did not include partners of affected men. We caution against generalizing the findings to the broader population of partners. The potential influence of selection bias is also an important consideration, as partners completing the assessment may differ from those who did not. The assessment used self-reported measures of communication channels, perceived need for psychoeducation, and other key constructs. Due to the anonymous, Internet-based nature of the needs assessment, it was also not possible to verify reported clinical characteristics (eg, the partner's receipt of *BRCA1/2* genetic testing or undergoing surgery).

Conclusions

These limitations notwithstanding, the findings highlight areas for future research with the goal of developing Internet-based psychoeducational interventions for partners of previvors and survivors of hereditary breast cancer. Though partners have diverse needs, use of self-help materials (either in print or electronic form) and interactive, Internet-based resources seems promising. Additional research is needed to rigorously develop and evaluate Internet-based resources targeting partners. Medical Internet and health communication researchers should work closely with cancer genetic content experts to examine ways in which Internet-based resources can be offered to this target population.

Acknowledgments

Preparation of this manuscript was supported in part by National Institutes of Health grants HG002686 and CA137625 to Dr Tercyak.

Conflicts of Interest

None declared.

References

1. Merrill RM, Sloan A. Risk-adjusted female breast cancer incidence rates in the United States. Cancer Epidemiol 2011 Sep 6. [doi: 10.1016/j.canep.2011.08.004] [Medline: 21903503]
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010 Dec 15;127(12):2893-2917. [doi: 10.1002/ijc.25516] [Medline: 21351269]
3. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A. Breast cancer before age 40 years. Semin Oncol 2009 Jun;36(3):237-249. [doi: 10.1053/j.seminoncol.2009.03.001] [Medline: 19460581]
4. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol 2007 Apr 10;25(11):1329-1333. [doi: 10.1200/JCO.2006.09.1066] [Medline: 17416853]
5. Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. J Natl Cancer Inst 2010 May 19;102(10):680-691 [FREE Full text] [doi: 10.1093/jnci/djq088] [Medline: 20427433]
6. Litton JK, Ready K, Chen H, Gutierrez-Barrera A, Etzel CJ, Meric-Bernstam F, et al. Earlier age of onset of BRCA mutation-related cancers in subsequent generations. Cancer 2012 Jan 15;118(2):321-325. [doi: 10.1002/cncr.26284] [Medline: 21913181]
7. National Comprehensive Cancer Network. 2011. NCCN Clinical Practice Guidelines in Oncology: Genetic/Familial High-Risk Assessment: Breast and Ovarian. Version 1.2011 URL: http://www.nccn.org/index.asp [accessed 2012-01-05] [WebCite Cache ID 64TRDBcrw]
8. National Comprehensive Cancer Network. 2010. NCCN Clinical Practice Guidelines in Oncology: Genetic/Familial High-Risk Assessment: Breast and Ovarian. Version 1.2010 URL: http://www.nccn.org/index.asp [accessed 2012-01-05] [WebCite Cache ID 64TRFyvNw]
9. Cheung EL, Olson AD, Yu TM, Han PZ, Beattie MS. Communication of BRCA results and family testing in 1,103 high-risk women. Cancer Epidemiol Biomarkers Prev 2010 Sep;19(9):2211-2219 [FREE Full text] [doi: 10.1158/1055-9965.EPI-10-0325] [Medline: 20699375]
10. Tercyak KP, Peshkin BN, DeMarco TA, Brogan BM, Lerman C. Parent-child factors and their effect on communicating BRCA1/2 test results to children. Patient Educ Couns 2002 Jun;47(2):145-153. [Medline: 12191538]
11. Tercyak KP, Peshkin BN, Streisand R, Lerman C. Psychological issues among children of hereditary breast cancer gene (BRCA1/2) testing participants. Psychooncology 2001;10(4):336-346. [Medline: 11462232]
12. Peshkin BN, DeMarco TA, Brogan BM, Lerman C, Isaacs C. BRCA1/2 testing: complex themes in result interpretation. J Clin Oncol 2001 May 1;19(9):2555-2565. [Medline: 11331335]
13. Robson ME, Storm CD, Weitzel J, Wollins DS, Offit K, American Society of Clinical Oncology. American Society of Clinical Oncology policy statement update: genetic and genomic testing for cancer susceptibility. J Clin Oncol 2010 Feb 10;28(5):893-901. [doi: 10.1200/JCO.2009.27.0660] [Medline: 20065170]
14. Bradbury AR, Dignam JJ, Ibe CN, Auh SL, Hlubocky FJ, Cummings SA, et al. How often do BRCA mutation carriers tell their young children of the family's risk for cancer? A study of parental disclosure of BRCA mutations to minors and young adults. J Clin Oncol 2007 Aug 20;25(24):3705-3711. [doi: 10.1200/JCO.2006.09.1900] [Medline: 17704419]

15. Patenaude AF, Dorval M, DiGianni LS, Schneider KA, Chittenden A, Garber JE. Sharing BRCA1/2 test results with first-degree relatives: factors predicting who women tell. J Clin Oncol 2006 Feb 1;24(4):700-706. [doi: 10.1200/JCO.2005.01.7541] [Medline: 16446344]
16. Demarco TA, Nusbaum RH, Peshkin BN, Patenaude AF, Schneider KA, Garber JE, et al. Prevalence and correlates of mothers and fathers attending pretest cancer genetic counseling together. Patient Educ Couns 2010 Jan;78(1):29-33. [doi: 10.1016/j.pec.2009.05.013] [Medline: 19545972]
17. Manne S, Audrain J, Schwartz M, Main D, Finch C, Lerman C. Associations between relationship support and psychological reactions of participants and partners to BRCA1 and BRCA2 testing in a clinic-based sample. Ann Behav Med 2004 Dec;28(3):211-225. [doi: 10.1207/s15324796abm2803_10] [Medline: 15576260]
18. Tercyak KP, Hughes C, Main D, Snyder C, Lynch JF, Lynch HT, et al. Parental communication of BRCA1/2 genetic test results to children. Patient Educ Couns 2001 Mar;42(3):213-224. [Medline: 11164320]
19. DeMarco TA, Peshkin BN, Valdimarsdottir HB, Patenaude AF, Schneider KA, Tercyak KP. Role of parenting relationship quality in communicating about maternal BRCA1/2 genetic test results with children. J Genet Couns 2008 Jun;17(3):283-287. [doi: 10.1007/s10897-007-9147-7] [Medline: 18288593]
20. Tercyak KP, Peshkin BN, Demarco TA, Patenaude AF, Schneider KA, Garber JE, et al. Information needs of mothers regarding communicating BRCA1/2 cancer genetic test results to their children. Genet Test 2007;11(3):249-255. [doi: 10.1089/gte.2006.0534] [Medline: 17949286]
21. Peshkin BN, Demarco TA, Tercyak KP. On the development of a decision support intervention for mothers undergoing BRCA1/2 cancer genetic testing regarding communicating test results to their children. Fam Cancer 2010 Mar;9(1):89-97. [doi: 10.1007/s10689-009-9267-3] [Medline: 19609726]
22. Budin WC, Hoskins CN, Haber J, Sherman DW, Maislin G, Cater JR, et al. Breast cancer: education, counseling, and adjustment among patients and partners: a randomized clinical trial. Nurs Res 2008;57(3):199-213. [doi: 10.1097/01.NNR.0000319496.67369.37] [Medline: 18496106]
23. Barsevick AM, Sweeney C, Haney E, Chung E. A systematic qualitative analysis of psychoeducational interventions for depression in patients with cancer. Oncol Nurs Forum 2002;29(1):73-84; quiz 85. [doi: 10.1188/02.ONF.73-87] [Medline: 11817494]
24. Devine EC, Westlake SK. The effects of psychoeducational care provided to adults with cancer: meta-analysis of 116 studies. Oncol Nurs Forum 1995 Oct;22(9):1369-1381. [Medline: 8539178]
25. Brotto LA, Heiman JR, Goff B, Greer B, Lentz GM, Swisher E, et al. A psychoeducational intervention for sexual dysfunction in women with gynecologic cancer. Arch Sex Behav 2008 Apr;37(2):317-329. [doi: 10.1007/s10508-007-9196-x] [Medline: 17680353]
26. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624 [FREE Full text] [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
27. Rutten LJ, Squiers L, Hesse B. Cancer-related information seeking: hints from the 2003 Health Information National Trends Survey (HINTS). J Health Commun 2006;11 Suppl 1:147-156. [doi: 10.1080/10810730600637574] [Medline: 16641080]
28. Bender JL, O'Grady L, Jadad AR. Supporting cancer patients through the continuum of care: a view from the age of social networks and computer-mediated communication. Curr Oncol 2008 Aug;15 Suppl 2:s107.es42-s107.es47 [FREE Full text] [Medline: 18769582]
29. 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]
30. Nguyen HQ, Carrieri-Kohlman V, Rankin SH, Slaughter R, Stulbarg MS. Internet-based patient education and support interventions: a review of evaluation studies and directions for future research. Comput Biol Med 2004 Mar;34(2):95-112. [doi: 10.1016/S0010-4825(03)00046-5] [Medline: 14972630]
31. Klemm P, Bunnell D, Cullen M, Soneji R, Gibbons P, Holecek A. Online cancer support groups: a review of the research literature. Comput Inform Nurs 2003;21(3):136-142. [Medline: 12792194]
32. Seale C. Gender accommodation in online cancer support groups. Health (London) 2006 Jul;10(3):345-360. [doi: 10.1177/1363459306064495] [Medline: 16775019]
33. Broom A. The eMale: prostate cancer, masculinity and online support as a challenge to medical expertise. J Sociol 2005;41(1):87-104. [doi: 10.1177/1440783305050965]
34. Cochrane BB, Lewis FM. Partner's adjustment to breast cancer: a critical analysis of intervention studies. Health Psychol 2005 May;24(3):327-332. [doi: 10.1037/0278-6133.24.3.327] [Medline: 15898870]
35. Bartholomew LK, Parcel GS, Kok G, Gottleib NH. Planning Health Promotion Programs: An Intervention Mapping Approach. 2nd edition. San Francisco, CA: Jossey-Bass; 2006.
36. Green LW, Kreuter MW. Health Promotion Planning: An Educational and Ecological Approach. 4th edition. New York, NY: McGraw-Hill; 2005.

37. Tercyak KP, Hensley Alford S, Emmons KM, Lipkus IM, Wilfond BS, McBride CM. Parents' attitudes toward pediatric genetic testing for common disease risk. Pediatrics 2011 May;127(5):e1288-e1295. [doi: 10.1542/peds.2010-0938] [Medline: 21502235]
38. Kumar V. Cluster analysis: basic concepts and algorithms. In: Tan PN, Steinbach M, Kumar V, editors. Introduction to Data Mining. Boston, MA: Pearson Addison Wesley; 2006:487-568.
39. Milligan GW, Cooper MC. Methodology review: clustering methods. Appl Psychol Meas 1987;11(4):329-354. [doi: 10.1177/014662168701100401]
40. Tabachnick BG, Fidell LS. Using Multivariate Statistics. 5th edition. Boston, MA: Pearson/Allyn & Bacon; 2007.
41. Little RJA. A test of missing completely at random for multivariate data with missing values. J Am Stat Assoc 1988;83(404):1198-1202.
42. Im EO, Chee W, Guevara E, Lim HJ, Liu Y, Shin H. Gender and ethnic differences in cancer patients' needs for help: an Internet survey. Int J Nurs Stud 2008 Aug;45(8):1192-1204. [doi: 10.1016/j.ijnurstu.2007.09.006] [Medline: 17963769]
43. Grande GE, Myers LB, Sutton SR. How do patients who participate in cancer support groups differ from those who do not? Psychooncology 2006 Apr;15(4):321-334. [doi: 10.1002/pon.956] [Medline: 16106472]
44. Mo PK, Malik SH, Coulson NS. Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups. Patient Educ Couns 2009 Apr;75(1):16-24. [doi: 10.1016/j.pec.2008.08.029] [Medline: 19233586]
45. Lieberman MA. Gender and online cancer support groups: issues facing male cancer patients. J Cancer Educ 2008;23(3):167-171. [doi: 10.1080/08858190802039185] [Medline: 18709588]
46. Klemm P, Hurst M, Dearholt SL, Trone SR. Gender differences on Internet cancer support groups. Comput Nurs 1999;17(2):65-72. [Medline: 10194883]
47. Zahlis EH, Lewis FM. Coming to grips with breast cancer: the spouse's experience with his wife's first six months. J Psychosoc Oncol 2010 Jan;28(1):79-97. [doi: 10.1080/07347330903438974] [Medline: 20391067]
48. Lewis FM, Fletcher KA, Cochrane BB, Fann JR. Predictors of depressed mood in spouses of women with breast cancer. J Clin Oncol 2008 Mar 10;26(8):1289-1295. [doi: 10.1200/JCO.2007.12.7159] [Medline: 18323552]
49. Lewis FM, Cochrane BB, Fletcher KA, Zahlis EH, Shands ME, Gralow JR, et al. Helping Her Heal: a pilot study of an educational counseling intervention for spouses of women with breast cancer. Psychooncology 2008 Feb;17(2):131-137. [doi: 10.1002/pon.1203] [Medline: 17429834]
50. Hordern A. Intimacy and sexuality for the woman with breast cancer. Cancer Nurs 2000 Jun;23(3):230-236. [Medline: 10851774]
51. Manne S, Ostroff J, Winkel G, Goldstein L, Fox K, Grana G. Posttraumatic growth after breast cancer: patient, partner, and couple perspectives. Psychosom Med 2004;66(3):442-454 [FREE Full text] [Medline: 15184709]
52. Fentiman IS, Fourquet A, Hortobagyi GN. Male breast cancer. Lancet 2006 Feb 18;367(9510):595-604. [doi: 10.1016/S0140-6736(06)68226-3] [Medline: 16488803]
53. Gómez-Raposo C, Zambrana Tévar F, Sereno Moyano M, López Gómez M, Casado E. Male breast cancer. Cancer Treat Rev 2010 Oct;36(6):451-457. [doi: 10.1016/j.ctrv.2010.02.002] [Medline: 20193984]
54. Liede A, Karlan BY, Narod SA. Cancer risks for male carriers of germline mutations in BRCA1 or BRCA2: a review of the literature. J Clin Oncol 2004 Feb 15;22(4):735-742. [doi: 10.1200/JCO.2004.05.055] [Medline: 14966099]
55. Hallowell N, Ardern-Jones A, Eeles R, Foster C, Lucassen A, Moynihan C, et al. Men's decision-making about predictive BRCA1/2 testing: the role of family. J Genet Couns 2005 Jun;14(3):207-217. [doi: 10.1007/s10897-005-0384-3] [Medline: 15959652]
56. 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]
57. Baker TB, Hawkins R, Pingree S, Roberts LJ, McDowell HE, Shaw BR, et al. Optimizing eHealth breast cancer interventions: which types of eHealth services are effective? Transl Behav Med 2011 Mar;1(1):134-145. [Medline: 21709810]
58. Gustafson DH, Hawkins R, McTavish F, Pingree S, Chen WC, Volrathongchai K, et al. Internet-based interactive support for cancer patients: are integrated systems better? J Commun 2008 Jun;58(2):238-257. [doi: 10.1111/j.1460-2466.2008.00383.x] [Medline: 21804645]
59. Marcus AC, Heimendinger J, Berman E, Strecher V, Bright MA, Allen AR, et al. A case study in dissemination: lessons learned from a pilot study involving the National Cancer Institute's Cancer Information Service. J Health Commun 2005;10 Suppl 1:219-233. [doi: 10.1080/10810730500257689] [Medline: 16377609]
60. Hawkins RB, Pingree S, Baker TB, Roberts LJ, Shaw BR, McDowell H. Integrating eHealth with human services for breast cancer patients. Transl Behav Med 2011;1(1):146-154. [doi: 10.1007/s13142-011-0027-1]
61. Alexander J, Zeibland S. The Web--bringing support and health information into the home: the communicative power of qualitative research. Int J Nurs Stud 2006 May;43(4):389-391. [doi: 10.1016/j.ijnurstu.2005.10.012] [Medline: 16438973]
62. Galinsky MJ, Schopler JH, Abell MD. Connecting group members through telephone and computer groups. Health Soc Work 1997 Aug;22(3):181-188. [Medline: 9260082]
63. Madara EJ. The mutual-aid self-help online revolution. Soc Policy 1997;27(3):20-26.

Abbreviations

FORCE: Facing Our Risk of Cancer Empowered
MANOVA: multivariate analysis of variance

Edited by G Eysenbach; submitted 20.05.11; peer-reviewed by E Matloff; comments to author 26.09.11; revised version received 17.10.11; accepted 18.10.11; published 18.01.12.

Please cite as:
Tercyak KP, Mays D, DeMarco TA, Sharff ME, Friedman S
Results of an Online Community Needs Assessment for Psychoeducational Interventions Among Partners of Hereditary Breast Cancer Previvors and Survivors
J Med Internet Res 2012;14(1):e15
URL: http://www.jmir.org/2012/1/e15/
doi:10.2196/jmir.1847
PMID:22257650

Original Paper

Comparing the Use of an Online Expert Health Network against Common Information Sources to Answer Health Questions

Martijn DF Rhebergen[1], MSc; Annet F Lenderink[2], MD; Frank JH van Dijk[1], MD, PhD; Carel TJ Hulshof[1,3], MD, PhD

[1]Academic Medical Center Amsterdam, Coronel Institute of Occupational Health, University of Amsterdam, Amsterdam, Netherlands

[2]Academic Medical Center Amsterdam, Dutch Center for Occupational Diseases, University of Amsterdam, Amsterdam, Netherlands

[3]Netherlands Society of Occupational Medicine (NVAB), Centre of Excellence, Utrecht, Netherlands

Corresponding Author:
Martijn DF Rhebergen, MSc
Academic Medical Center Amsterdam
Coronel Institute of Occupational Health
University of Amsterdam
Meibergdreef 9
P.O. Box: 22700
Amsterdam, 1100 DE
Netherlands
Phone: 31 20566 ext 2831
Fax: 31 20 697 7161
Email: m.d.rhebergen@amc.uva.nl

Abstract

Background: Many workers have questions about occupational safety and health (OSH). It is unknown whether workers are able to find correct, evidence-based answers to OSH questions when they use common information sources, such as websites, or whether they would benefit from using an easily accessible, free-of-charge online network of OSH experts providing advice.

Objective: To assess the rate of correct, evidence-based answers to OSH questions in a group of workers who used an online network of OSH experts (intervention group) compared with a group of workers who used common information sources (control group).

Methods: In a quasi-experimental study, workers in the intervention and control groups were randomly offered 2 questions from a pool of 16 standardized OSH questions. Both questions were sent by mail to all participants, who had 3 weeks to answer them. The intervention group was instructed to use only the online network ArboAntwoord, a network of about 80 OSH experts, to solve the questions. The control group was instructed that they could use all information sources available to them. To assess answer correctness as the main study outcome, 16 standardized correct model answers were constructed with the help of reviewers who performed literature searches. Subsequently, the answers provided by all participants in the intervention (n = 94 answers) and control groups (n = 124 answers) were blinded and compared with the correct model answers on the degree of correctness.

Results: Of the 94 answers given by participants in the intervention group, 58 were correct (62%), compared with 24 of the 124 answers (19%) in the control group, who mainly used informational websites found via Google. The difference between the 2 groups was significant (rate difference = 43%, 95% confidence interval [CI] 30%–54%). Additional analysis showed that the rate of correct main conclusions of the answers was 85 of 94 answers (90%) in the intervention group and 75 of 124 answers (61%) in the control group (rate difference = 29%, 95% CI 19%–40%). Remarkably, we could not identify differences between workers who provided correct answers and workers who did not on how they experienced the credibility, completeness, and applicability of the information found ($P > .05$).

Conclusions: Workers are often unable to find correct answers to OSH questions when using common information sources, generally informational websites. Because workers frequently misjudge the quality of the information they find, other strategies are required to assist workers in finding correct answers. Expert advice provided through an online expert network can be effective for this purpose. As many people experience difficulties in finding correct answers to their health questions, expert networks may be an attractive new source of information for health fields in general.

(J Med Internet Res 2012;14(1):e9) doi:10.2196/jmir.1886

KEYWORDS

Information services; online expert network; medical informatics; information-seeking behavior; occupational health; evidence-based practice; question and answer

Introduction

Many workers seek information to answer occupational safety and health (OSH) questions [1-4]. For this purpose, workers commonly use a variety of sources, such as asking advice from OSH experts within their company or social network, or exploring informational websites [5-7]. Ideally, the information available from these common information sources is of high quality [6,7], as low-quality information may lead to incorrect answers and wrong decisions regarding the prevention or management of OSH at work [8]. It is unknown whether workers can find correct, evidence-based answers [9,10].

In general, one might ask whether it is possible for workers to find correct answers to their OSH questions. Finding information and answering health-related questions require specific skills or literacy [11,12]. The World Health Organization defines health literacy as "the cognitive and social skills which determine the motivation and the ability of an individual to gain access to, understand and use information in ways that promote and maintain good health" [13]. The Internet, the source that workers use most frequently, often provides excessive amounts of information that is not always easy to understand or of high quality [14-17]. The solution might be to consult OSH experts who are trained in finding evidence-based answers to clinical questions [18,19]. However, the consultation of experts by workers might be hampered by restricted access and high costs [20]. An attractive solution might be a selection of the best of both options by offering workers easily accessible, free-of-charge online advice from OSH experts [21-23].

Question-and-answer expert network tools could be useful for providing such online expert advice, as these tools are designed for communication, and knowledge dissemination, storage, and retrieval [22,23]. These tools have the potential to build a network of experts on a particular topic (and many subtopics) and make them accessible to questioners. We tested the hypothesis that workers who used this online network of OSH experts would find correct answers to OSH questions more often than workers who used commonly available information sources. The aim of this study was to answer the following question: is there a difference in the rate of correct answers to questions about OSH between workers who use expert advice through an online expert network and workers who use common information sources?

Methods

Study Design

In a quasi-experimental study, we compared the rate of correct answers formulated by a group of workers who used an online expert network with that of a group of workers who used common information sources.

Online Network on OSH: ArboAntwoord

ArboAntwoord is an experimental, free-of-charge facility for workers with OSH questions [23,24]. The network was launched by means of a small-scale campaign in October 1, 2008. The home page of ArboAntwoord comprises several main categories of leading OSH topics (Figure 1). After registration, a worker can pose his or her question directly in the designated text field on the home page, or he or she can use the button "ask your question" that is presented in all subcategories. Both possibilities lead to a webpage in which the question must be given a title and the questioner must prohibit or authorize the publication of the question (Figure 2). After formulating his or her question, the worker needs to select one or more experts in that subcategory (Figure 3). Every expert is indexed to the subcategory that corresponds to his or her expertise. Appreciation scores expressed by earlier questioners and mean reaction time to previously answered questions are provided to facilitate questioners' choosing an expert. Questioners may choose more than one expert. To notify the selected expert about his or her asked question, the page provides a "send question to the expert" button. With this button the selected expert will receive an email notification with a direct hyperlink to the question. Here the expert is also provided a main text field where he or she can also add an attachment when wanted (Figure 4). An expert can notify the questioner about his or her answer by using a "send answer to the questioner" button. All stored questions and answers are published and can be searched by other users when authorized by the questioner and the moderator. When required, experts can react to published questions and answers. Eligible questions and answers are stored and made accessible to other questioners in a searchable database after moderation and after informed consent by the questioner (Figure 5).

A steady network of about 80 experts participate in the network. All experts are invited and/or accepted to participate if they meet all of the following criteria: (1) working in a university, or a commercial OSH expert center or OSH organization operating on a national level, (2) having (inter)national expertise on a specific OSH topic, (3) having at least 5 years of experience on this topic, and (4) participating in at least one knowledge-dissemination activity such as authorship of scientific articles or participation in an expert committee. The professions of the experts vary: occupational physicians, hygienists, safety workers, health scientists, psychologists, neuropsychologists, and experts in OSH law and regulations. All ArboAntwoord experts participate on personal title and on voluntarily basis. Discussion among experts is not common, although about a third of the questions are answered by more than one expert.

Figure 1. Screenshot of the ArboAntwoord home page, where a questioner can select a question category and use a search function to find stored questions and answers.

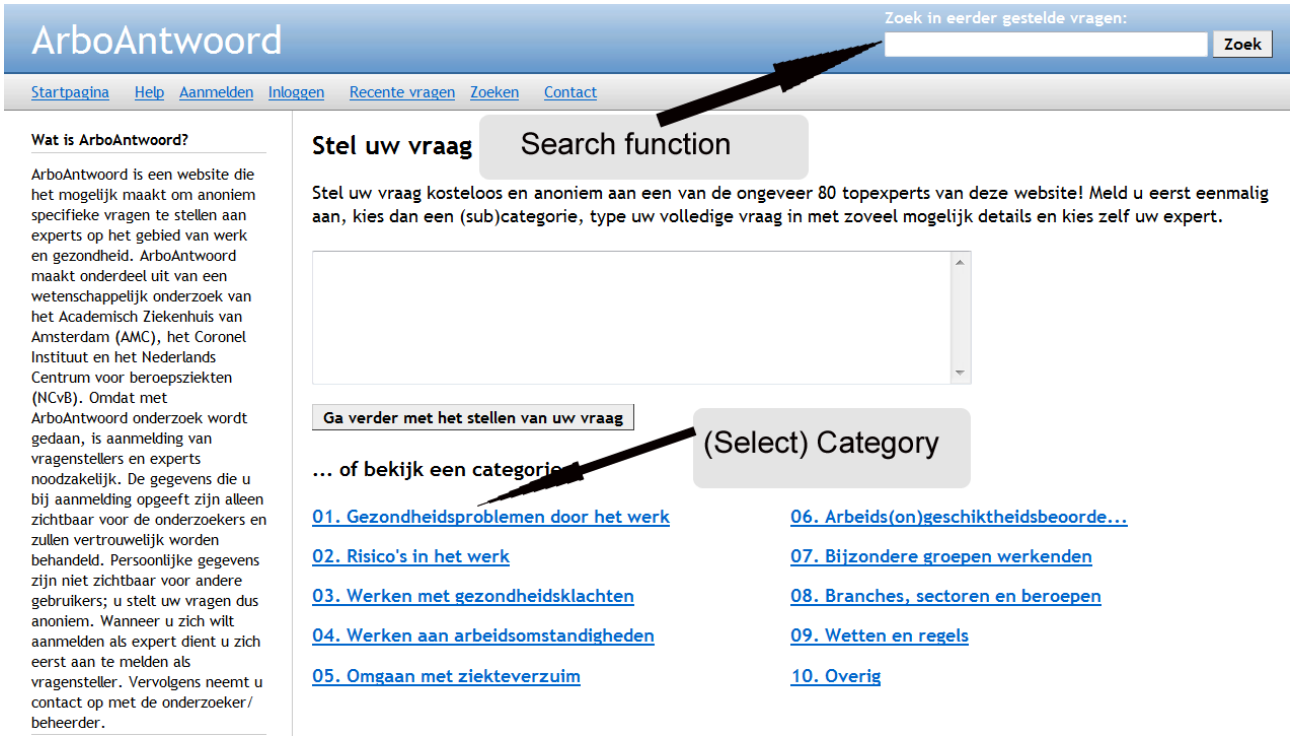
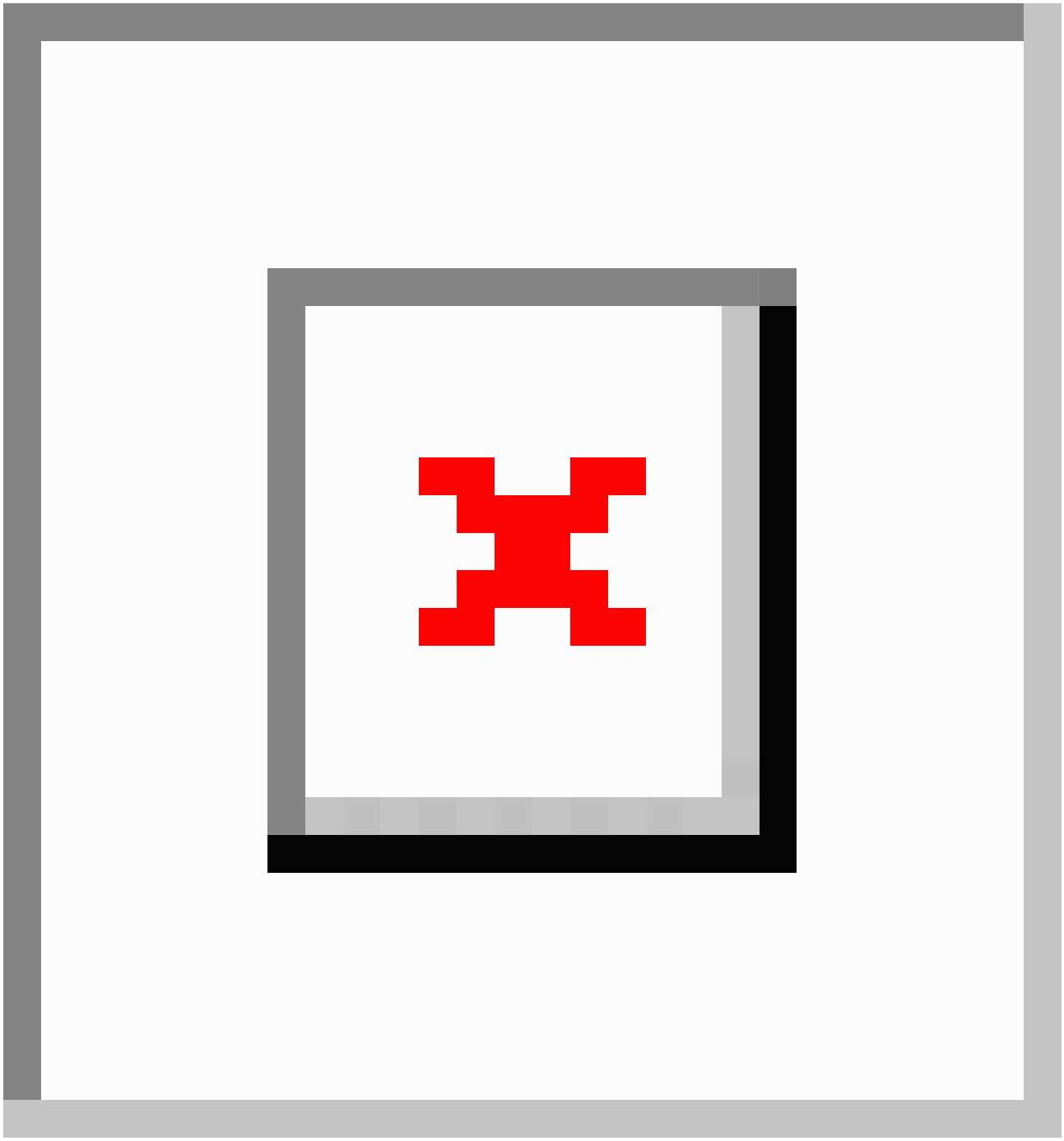


Figure 2. Screenshot of the webpage in a subcategory, where workers actually pose their question, give the question a title, and authorize publication.

Figure 3. Screenshot of the webpage in a subcategory, where a questioner can select one or more experts. The webpage also includes the experts' mean reaction time and appreciation and a button to send the question.

Figure 4. Screenshot of the webpage in a subcategory, where an expert can provide his or her answer, with or without an attachment. The page includes a button to send the question.

Figure 5. Screenshot of the webpage in a subcategory that provides workers with a hyperlink to view stored and recently asked questions and answers in that subcategory.



Intervention Group

The intervention group was instructed to use the online network ArboAntwoord for solving 2 OSH questions that were provided by the researchers. As we did not have the opportunity to observe and log their use of the online network directly, we incorporated the 2 OSH questions into paper logs. These logs were mailed to the participants. A log included a question, a field to provide an answer, and several secondary outcome-related questions. A short additional questionnaire for information on background characteristics such as gender, age, educational level, work role, company size, and company sector was sent with the logs. We mentioned that ArboAntwoord was an experimental, free-of-charge online network of experts answering workers' OSH questions, and we explained that they had to register on the website. Participants were instructed to pose a question in the (subcategory) they thought was the most convenient. The participants were requested to provide a clear and complete answer. Lastly, participants were asked to return the completed logs to the researchers in the return envelop provided. They could also deliver the logs with the start of the course, explained below. Participants had a maximum of 3 weeks to fill in the logs. As an incentive, the participants were promised a memory stick with useful OSH information.

Control Group

The control group can be interpreted as a care-as-usual group, using the information-seeking strategies and sources of workers

with OSH questions in daily practice. Similar to the intervention group, the control group was sent 2 OSH questions that were incorporated into paper logs. These logs included a question, a space to provide the answer, and several secondary outcome-related questions. The logs of the controls instructed them that "You may use all (types of) information sources available to you to solve the questions." Additionally, we asked them where they looked for information. For this question we provided 4 answer categories: (1) Internet using Google, (2) written sources (ie, magazines and books), (3) experts or specialists, and (4) other sources of information. In addition to this question, controls were requested to note the source that provided the most relevant information. Lastly, we asked the question "How much time in minutes did you spent seeking the information?" The participants were requested to provide a clear and complete answer. They were asked to fill in the logs and return these to the researchers. In total, participants got about 3 weeks to finish the logs. As an incentive the participants were promised a memory stick with useful OSH information.

Participants

Because we needed a motivated group of workers for this study, we decided to recruit workers who were enrolled to take part in a course for OSH supervisor (in Dutch: preventiemedewerker). An OSH supervisor is a common worker who is responsible for recognizing and suggesting basic solutions for OSH risks in a company. In general, short courses lasting 1–2 days on basic OSH issues are provided to educate and train new supervisors. Although workers enrolled in these courses probably have a higher than average interest in OSH, we know from training reports that more than 80% of the workers start without any substantial knowledge of OSH. Therefore, before the start of a course, we consider enrolled workers to be OSH interested, but not OSH educated. We approached all enlisted course participants from 2 training organizations during 2010 (N = 192). There were no important differences in the content of the course programs. Workers enrolled in the OSH supervisor courses prior to July 2010 were allocated to a control group (courses 1–16; n = 105). All workers enrolled in a course starting in August 2010 or later that year were assigned to an intervention group (courses 17–29; n = 87).

Data Collection

Both groups were offered 2 questions from a pool of 16 standardized OSH questions. As the difficulty and the topic of the question may affect the likelihood of finding a correct answer, we included these aspects in the construction of the question pool. For question difficulty, we distinguished *simple* information or knowledge questions and *complex* interpretation or advice questions. A simple question was defined as an OSH question that could be answered directly by one specific piece of information or advice. A complex question was defined as a question that could only be answered by interpreting and combining several pieces of information, often accounting for contextual aspects. We formulated 8 simple and 8 complex questions. We further distinguished 2 questions topics. Of the 16 questions, 5 considered OSH laws and regulations (eg, Are safety shoes obligatory for workers in an army storage depot?). The remaining 11 questions were about actual OSH issues, such as causes and risks factors for work-related health and safety (eg, Is radiation a risk for pregnant magnetic resonance imaging workers?), diagnosis (eg, What are the diagnostic criteria for designating posttraumatic stress disorder as an occupational disease?), interventions in occupational safety or health (eg, What is an effective intervention for occupational dermatitis?), or social mapping (eg, Where can I find the best expert on chronic low back pain rehabilitation?).

To increase the ecological validity (the extent to which research emulates the real world) of the 16 OSH questions to be selected, we randomly selected 12 questions with answers from the ArboAntwoord database [25]. Additionally, 4 simple questions were formulated by the researchers based on information provided in the OSH supervisor course handbooks, and these questions were added to the pool. For the random selection of the 12 ArboAntwoord questions, all 319 questions in the ArboAntwoord database were first stratified by 2 researchers independently (MR and AF) in 4 categories based on difficulty (simple or complex) and topic (OSH law and regulations, or OSH content). Subsequently, 2 simple OSH law and regulation questions, 3 complex OSH law and regulation questions, 6 simple OSH content questions, and 5 complex OSH content questions were used in this study. An example of a difficult OSH content question is "Are the glass fibers or dust released after the crushing, cutting or fragmentation of (car) windows in the open air hazardous to my health? What can be done to prevent hazards?" We observed that the structure of 8 questions included 2 components—that is, these questions were actually composed of 2 questions. This corresponds to our experience with questions from practice, where workers often seem to have a concern about an OSH risk and wonder about a possible solution. The 16 questions are included in Multimedia Appendix 1.

We randomly assigned 2 questions to each participant and sent them out about 3 weeks before the start of their OSH supervisor course, ensuring that all participants received 1 easy and 1 complex question (to ensure they would not be discouraged by getting 2 complex questions). Based on the level of question difficulty, an automatically created randomization list was used with 56 possible combinations (8 simple × 8 – 1 complex questions) for both the intervention and control groups. Because we anticipated about 200 course participants and a response rate of 50%, we expected that every question would be answered about 6 times in the control group and 6 times in the intervention group. Although participants in the intervention group posed a question directly to an ArboAntwoord expert, the system moderator (MR) always provided questioners with the original answers to 12 randomly selected OSH questions to prevent unnecessary use of expert time and to avert possible learning effects resulting from answering the same question more than once. For the 4 self-formulated OSH questions we used the first answer provided by an expert in our experiment.

Primary Outcome and Scoring Procedure

To define the main study outcome parameter *answer correctness* we used the experiences and approaches of evidence-based practice [9,10]. Evidence-based practice is a strategy for clinical decision making that involves the integration of the patients'

needs and context, the expertise of professionals, and the best available research evidence. In this study, we define a correct answer as "an answer that accounts for the context of the question(er) and corresponds with conclusions or recommendations of the best available evidence". Two steps were needed to assess whether provided answers were correct.

First, 16 standardized correct *model* answers were constructed with the help of 2 reviewers with expertise on the question topics. A total of 10 reviewers were internal experts who also participated in the ArboAntwoord network, and 13 external reviewers were recruited for the review process. Internal experts could not review their own answers. All reviewers were provided with a draft model answer formulated by the research team. The reviewers were requested to report whether the draft model answer was in complete, partial, or no agreement with the conclusions or recommendations from the best available evidence. Similarly to the evidence-based practice method, reviewers were asked to consider 4 levels of evidence by means of an evidence literature search looking for evidence-based guidelines, reviews, or scientific research articles, or using their own professional expertise, in that order. If evidence from the highest level was not available, we requested the expert to move on to the next best level of available evidence. Because these levels of evidence do not apply to OSH law and regulation questions, reviewers of draft answers on this topic were requested to first regard laws and regulations, policy, jurisdictions, or standards, followed by their own professional expertise. All additional evidence and comments provided by both reviewers were included in the final versions of the standardized correct model answers.

Second, we developed a scoring system defining correctness for all 16 standardized model answers separately based on the essential aspects required for a correct answer (see Multimedia Appendix 1). The scoring system consisted of an answer correctness score between 0 and 4. The essential aspects were directly dependent on the nature of the particular OSH question. For example, questions sometimes asked for confirmation or proof, for conditions under which an OSH situation holds, for a specific location, for possible risks or solutions, or for a combination of these. To verify whether a participant finally reached the correct answer conclusion to the main question, we included this as one of the essential aspects in all 16 model answers. This aspect can be considered as the correctness of an answer's main conclusion (yes, no, or possibly). The question about the health hazards of glass fibers and dust discussed earlier in the Methods was scored as follows. The main conclusion, "Yes, this could possibly be hazardous to health," was given 1 point. Another point was provided when the answer mentioned something like "Depending on glass (particle) type and exposure." Two more points were awarded when a security measure were given. As in health care in general, there is sometimes some variation or interpretation for what constitutes a correct, evidence-based answer. This especially holds true for the questions related to OSH content. For all questions some room was created to account for this variation and interpretation issue (Multimedia Appendix 1). Based on the scoring system, all participants' answers were scored and compared with the model answers by 2 raters. The first rater was a medical student unrelated to this study. The second rater, MR, checked the answer scores of the first rater. Both raters were blinded to the group (intervention or control) to which the answers belonged.

Secondary Outcomes

We assessed 2 secondary outcomes: (1) the experienced quality of the information source used: whether the source was usable or easy to use, and how easy it was to learn to use, and (2) the experienced quality of the information obtained: whether the information was complete, applicable, and reliable, and how satisfied the participant was with the information. All response categories to these questions were based on 7-point Likert scales (ranging from completely disagree to completely agree).

Data Analysis

We described most outcomes by means of descriptive analysis. Analyses were performed with SPSS version 17.0 (IBM Corporation, Somers, NY, USA). To establish any group differences in background characteristics that required adjustment in further analysis, we first applied the chi-square test for dichotomous or nominal variables, and the Yates and Cochrane test for ordinal variables ($P < .05$).

The correctness of the participants' answers was analyzed in 2 ways. First, we verified whether the answers given were sufficiently correct by dichotomizing the 4-point answer scores (0–2 points = insufficiently evidence based; 3–4 = sufficiently evidence based). We considered using the ordinal data, but we observed that the distribution was skewed. Thus, dichotomization seemed the best option without giving away a lot of information. Second, we looked at the correct main conclusions of the answers. Possible group differences between the intervention and control groups regarding the prevalence of correct answers and of correct main conclusions were analyzed with the chi-square test ($P < .05$). Because we found no differences between the groups in terms of question difficulty, question topic, or background characteristics ($P > .05$), we used binary logistic regression analysis only to establish possible interaction effects of these factors on the main outcome(s) ($P < .05$). We stratified the effect of group type on the number of correct answers by question difficulty, question topic, question structure, and background characteristics. We used the chi-square test for dichotomous and nominal variables, and the Yates and Cochrane test for the ordinal background characteristics ($P < .05$). To describe the strength of associations, we used rate differences. In this study, rate difference is probably the most appropriate measure because it describes the absolute change in the rate of, for example, correct answers attributable to the intervention.

Potential differences between the groups regarding the 2 secondary outcomes (the experienced quality of the information source used and the information it provided) were analyzed with the Wilcoxon rank sum test ($P < .05$). Because we observed moderate to high Spearman correlations ($r = .45–.65$) and good internal consistency (Cronbach alpha = .74) between the 4 items on the experienced quality of the source used, these 4 items were processed into a single-item factor by calculating mean scores. We observed no high correlations between the 4 items on information quality ($r < .40$).

In the control group we also applied the Wilcoxon rank sum test to determine the effect of information-seeking time and experienced information quality or quality of the information source used on providing correct answers.

Results

Group Characteristics

Overall, 47 of the 87 (54%) workers assigned to the intervention group agreed to participate in the study compared with 62 of the 105 (59%) in the control group. This resulted in 94 answers in the intervention and 124 answers in the control group. In total 110 of the 124 (89%) questions in the control group were answered with information obtained online, 9 (7%) with information from written sources (ie, magazines, books), and 5 (4%) with advice from experts or specialists. Because removing questions that were answered with written information or expert advice did not change any of the outcomes, these questions were preserved in further analysis. The median information-seeking time in the control group was 10 minutes per question (interquartile range: 5–20 minutes), and this time in minutes was not comparable with the time in days the intervention group had to wait for their answer of the ArboAntwoord experts. We did not observe any significant group differences in background characteristics ($P > .05$) (Table 1). Young participants between the age of 15 and 24 years were not represented in either group.

Table 1. Background characteristics of the intervention group (n = 47) and the control group (n = 62)

Characteristic	Intervention group (online network ArboAntwoord)		Control group (common information sources)	
	n	%	n	%
Gender				
Female	28	60	33	53
Male	19	40	29	47
Age group (years)				
15–24	0	0	0	0
25–34	17	36	14	23
35–44	12	25	22	35
45–54	13	28	20	32
≥55	5	11	6	10
Educational level				
Low	11	23	13	21
Intermediate	17	36	21	34
High	19	40	28	45
Role				
Worker	32	68	42	68
Employer/manager	8	17	13	21
OSH[a] (semi)professional	7	15	7	11
Company size				
Small	19	40	27	44
Medium	16	34	18	29
Large	12	26	17	27
Company s ector				
Agriculture and fishery	2	4	3	5
Industry	13	28	12	19
Construction industry	4	10	6	10
Trade	8	17	11	17
Transport and communication	1	2	3	5
Financial services	1	2	1	2
Business services	8	17	7	11
Public policy or civil service	3	6	3	5
Education	1	2	3	5
Health care	3	6	7	11
Culture and other services	3	6	6	10
Self-rated Internet and computer experience				
Relatively inexperienced	8	17	16	26
Relatively experienced	39	83	46	74

[a] Occupational safety and health.

We observed no statistical group differences in the distribution of simple or complex questions ($\chi^2_1 = 0$; $P = .9$) and OSH legislation or OSH content questions ($\chi^2_1 = 0$; $P = .9$). Nevertheless, we observed a discrepancy in the distribution of

the answers to 2 of the OSH questions. Question 7 was answered once in the intervention group and 4 times in the control group, and question 13 was not answered at all in the intervention group and was answered 8 times in the control group. Because removing these 2 questions did not change any of the outcomes, both were preserved in further analysis. All other questions were answered between 4 and 10 times in both groups.

Answer Correctness

In total, 58 of the 94 (62%) answers of the intervention group were rated correct, compared with 24 of the 124 (19%) answers for the control group. A significant difference with a rate difference of 43% was observed (95% CI 30%–54%) (Table 2). The use of the online expert network ArboAntwoord had a positive effect on providing correct answers, and the effect was identical for answers to simple or complex questions, for answers to questions related to OSH law and regulations or OSH content, and for answers to single or double questions (Table 2). Stratification by background characteristics consistently showed similar differences in the distribution of evidence-based answers in favor of the intervention group, with rate differences ranging from 33% to 66% (Table 2). Only Internet use significantly interacted with the effect of intervention on answer correctness: using the online network ArboAntwoord for providing correct answers was found to be even more beneficial for relatively inexperienced Internet and computer users (Wald $\chi^2_1 = 3.9$; $P = .048$).

In the intervention group, answers to questions about OSH law and regulations were significantly more often correct than questions about OSH content ($\chi^2_1 = 7.9$; $P = .01$). This same trend was observed in the control group ($\chi^2_1 = 3.2$; $P = .07$). Furthermore, within the control group we found that correct answers were significantly less often provided by men than by women ($\chi^2_1 = 5.7$; $P = .02$), with a similar trend for gender in the intervention group ($\chi^2_1 = 3.7$; $P = .06$). Spending more time seeking information did not affect the likelihood of finding a correct answer ($Z = –0.18$, $P = .9$).

Finally, we analyzed a subgroup on the one essential aspect that was similar for all answers: the correctness of the main conclusion (yes, no, or possibly). In total, 85 of the 94 (90%) main conclusions in the intervention group were correct compared with 75 of the 124 (61%) conclusions for the control group. A significant difference with a rate difference of 29% was found (95% CI 19%–40%) (Table 3). This positive effect in favor of the intervention was identical for answers to simple and complex questions, for answers to questions related to OSH law and regulations and OSH content, and for answers to single and double questions.

Table 2. Rates of correct answers in the intervention group (n = 94 answers) compared with the control group (n = 124 answers) stratified by question difficulty, question topic, question structure, and background characteristics

	Intervention group (online network ArboAntwoord)		Control group (common information sources)		Intervention vs control group	
	n/N	%	n/N	%	RD%[a]	95% CI[b]
Total (all questions)	58/94	62	24/124	19	43	30–54
Question difficulty						
Simple	25/47	53	12/63	19	34	16–50
Complex	33/47	70	12/61	20	50	33–65
Question topic						
OSH[c] law and regulations	24/29	83	11/38	29	54	31–71
OSH content	34/65	52	13/86	15	37	22–51
Question structure						
Single	18/31	58	10/50	20	38	17–57
Double	40/63	63	14/74	19	44	29–58
Gender						
Female	39/56	70	18/66	27	43	25–57
Male	19/38	50	6/58	10	40	22–56
Age group (years)						
15–24	NA[d]		NA		NA	
25–34	22/34	65	6/28	21	44	19–63
35–44	15/24	63	10/44	23	40	15–60
45–54	16/26	62	6/40	15	47	23–66
≥55	5/10	50	2/12	17	33	–7 to 66
Educational level						
Low	14/22	64	4/26	15	49	21–69
Intermediate	22/34	65	9/42	21	44	21–61
High	22/38	58	11/56	20	38	19–56
Role						
Worker	40/64	63	20/84	24	39	23–53
Employer/manager	7/16	44	1/26	4	40	16–64
OSH (semi)professional	11/14	79	3/14	21	58	20–80
Company size						
Small	22/38	58	9/54	17	41	22–58
Medium	16/32	50	5/36	14	36	14–55
Large	20/24	83	10/34	29	54	29–72
Self-rated Internet and computer experience						
Inexperienced	12/16	75	3/32	9	66	38–8
Experienced	46/78	59	21/92	23	36	22–49

[a] Rate difference.

[b] Confidence interval.

[c] Occupational safety and health.

[d] Not applicable.

Table 3. Rates of correct main conclusions of answers in the intervention group (n = 94 answers) compared with the control group (n = 124 answers) stratified by question difficulty topic and structure

	Intervention group (online network ArboAntwoord)		Control group (common information sources)		Intervention vs control group	
	n/N	%	n/N	%	RD%[a]	95% CI[b]
Total (all questions)	85/94	90	75/124	61	29	19–40
Question difficulty						
Simple	41/47	87	40/63	64	23	7–38
Complex	44/47	94	35/61	57	37	21–50
Question topic						
OSH[a] law and regulations	28/29	97	26/38	68	29	11–45
OSH content	57/65	88	49/86	57	31	17–43
Question structure						
Single	31/31	100	31/50	62	38	26–52
Double	54/63	86	44/74	60	26	11–40

[a] Rate difference.

[b] Confidence interval.

[c] Occupational safety and health.

Experienced Quality of the Information (Sources) Used

On average, the online network ArboAntwoord was rated of higher quality than common information sources, with mean scores of 5.8 (interquartile range: 5.5–6.3) and 5.2 (interquartile range: 4.4–6.0), respectively ($Z = –3.5$, $P < .001$). Participants in the intervention group experienced the completeness of the received information to be significantly higher (Z = –2.6, P = .01) and were significantly more satisfied with the received information (Z = –2.3, P = .03) than participants in the control group (Table 4). Notably, judging the quality of the information seemed to be difficult for the participants. Within both the intervention and the control group, we did not find a significant difference in the experienced information quality scores between workers who provided correct answers and those who did not (within the intervention group: information completeness $Z = –0.9$, $P = .4$, applicability $Z = –1.0$, $P = .3$, and credibility $Z = –1.5$, $P = .1$; within the control group: information completeness $Z = –0.8$, $P = .4$, applicability $Z = –1.3$, $P = .2$, and credibility $Z = –0.6$, $P = .5$). This finding corresponds to the comparably high applicability and credibility scores found for both ArboAntwoord and the common information sources.

Table 4. Comparison between the intervention group (n = 94 answers) and the control group (n = 124 answers) regarding experienced information completeness, applicability, credibility, and satisfaction with the information

Experienced information quality	Intervention group (online network ArboAntwoord)		Control group (common information sources)		Intervention vs control group *Z*	Intervention vs control group *P* value
	Mean	IQR[a]	Mean	IQR		
Completeness	5.4	5.0–6.0	4.7	3.0–6.0	–2.6	.01
Applicability	5.5	5.0–6.0	5.3	5.0–6.0	–1.2	.2
Credibility	5.4	4.0–6.0	5.4	5.0–6.0	–0.05	.9
Satisfaction	5.6	6.0–6.0	5.0	4.8–6.0	–2.3	.03

[a] Interquartile range.

Discussion

Principal Results

Our findings show that the rate of correct answers to OSH questions provided by workers who used expert advice obtained from an online network was significantly higher than the rate of correct answers provided by workers who used common information sources. When workers used their common information sources (in 90% of the cases, these were informational websites found through Google), only 19% of the answers were correct. The rate of correct answers was 62% for workers using the online expert network ArboAntwoord, which is significantly higher. This difference was found for answers to simple and complex questions, for answers to questions about OSH law and regulations and about OSH

content, and for answers to single and double questions. Answer correctness rates in both groups increased to 90% and 61% when we analyzed only the correctness of the main conclusion of the answers. Overall, workers who used ArboAntwoord were more satisfied with information that they received, and they experienced the information as more complete than workers who used common information sources. Nevertheless, the perceived information quality scores were relatively high in both groups. Remarkably, within both the experimental and control groups, workers who provided incorrect answers believed the information that they used to be as credible, complete, and applicable as did workers who provided correct answers. Workers appear to be unable to judge the quality of the information they find.

Comparison with Prior Work

To our knowledge, this is one of the first studies evaluating whether and how workers can find correct evidence-based answers to OSH questions. So far, most studies on answering OSH questions focus on OSH professionals and their use of evidence-based practice strategies [6,7,18,19,26]. Our findings clearly demonstrate that workers provide correct answers more often with the support of expert advice than with common information sources. Because we checked that all information necessary to answer the 16 questions correctly was available on the Web, our findings might be partly explained by the fact that workers have limited search skills and seem unable to judge important qualities of the information they find. Consequently, workers who wrongly judge the credibility, completeness, and applicability of information as high are likely to provide an incorrect answer. This especially holds true for participants in the control group, who often used online information found with Google. In several studies, non-health professionals have been found to use too few search terms and to select only one of the first few results displayed by a search engine [27-29]. Moreover, the quality of information found with these search engines has been shown to vary [14,16,30-32]. Consequently, using search engines such as Google does not always result in finding correct evidence-based answers. Thiele et al [33] demonstrated that medical students, resident physicians, and attending physicians provided only about 65% correct answers to 8 anesthesia and critical care-based clinical questions using Google. Similarly, Kingsley et al [34] found that only 25% of first-year dental students provided correct answers to several fundamental biomedical questions when using Google. In a study by Tang and Ng [17], several samples of diagnostic cases were selected from the *New England Journal of Medicine* that were subsequently googled for a diagnosis. The searches revealed a correct diagnosis in only 58% of the cases. Moreover, Kortum et al [35] showed that nonprofessionals (high school students) often provide incorrect answers to health questions when using the Internet. Consistently with our results, they concluded that difficulties in distinguishing trustworthy from untrustworthy medical information resulted in these incorrect answers. In sum, similar to the concept of health literacy [13], OSH literacy corresponding to "the cognitive and social skills which determine the motivation and the ability of an individual to gain access to, understand and use information in ways that promote and maintain good occupational safety and health" seems to influence workers' ability to find correct answers to OSH questions. It would be worthwhile to further explore OSH literacy and its role in finding correct answers effectively.

The higher rate of correct answers in the intervention group was probably further amplified by the high proficiency of the experts associated with ArboAntwoord. They are leading national experts on specific OSH topics and are familiar with finding, selecting, appraising, and applying evidence-based information. Nevertheless, even these experts did not always provide a correct answer, which is in accordance with the findings of Schaafsma et al [36], who stated that caution is required when relying blindly on expert advice. Again, it is possible that participants sometimes did not understand or interpret the information of the experts correctly. Another possible explanation is that experts were hindered by a lack of time in answering questions thoroughly and performing evidence searches when needed [26]. Moreover, the experts in ArboAntwoord participated voluntarily and were not paid. In commercial, nonhealth-related networks, (perceived) answer quality was shown to increase when users were paid to provide answers [37]. Possibly, to further increase answer correctness, experts could be given small incentives and more thorough instruction in how to answer questions correctly.

Subgroup analysis resulted in additional interesting findings. Within the intervention and control groups, we observed that the rate of correct main conclusions was much higher than the rate of correct answers in general. It is possible that workers are often able to provide a sort of "logical" conclusion based on deduction, observation of current practices, common sense about moral responsibility, or even implicit knowledge. Additionally, the a priori chance to provide a correct conclusion is 33% (yes, no, or possibly), which may also partly explain why, in both groups, the rate of correct conclusions was higher than the rate of correct answers in general.

In both the intervention and control groups, the rate of correct answers to questions about OSH law and regulations was higher than that of answers to questions about OSH content. Apparently, OSH content questions are more difficult to answer than questions about OSH legislation for both questioners and experts. Possibly, the (poor) formulation of the OSH content questions might have made them more difficult to answer. Finally, in both the control and intervention groups, we found a trend that women seemed to outperform men. It is possible that men feel less obligated, motivated, or aroused to find correct answers to the questions, or they may have less efficient learning styles (including information-seeking strategies) that affect the effort of seeking information [38,39].

Two expected effects could not be established in our analyses. We expected that for complex questions the rate of correct answers would be significantly lower than for simple questions, especially in the control group. We presumed that expert advice would be particularly necessary for the complex questions. Our findings did not corroborate these expectations. Our hypothesis may have been incorrect, or it is possible that our simple questions were not actually very simple, or our complex questions were not actually very complex. Finally, we expected that information-seeking time would influence the rate of correct

answers in the control group: that spending more time seeking information would have a positive influence on this rate. However, this difference could not be established. Again, because workers often seem to misjudge information quality, they believe that spending more time on information seeking is unnecessary. Another explanation might be that self-reporting of the time spent seeking information is subject to social desirability bias. Estimating the time required might therefore be less reliable than actually observing and timing the information-seeking process.

Strengths and Limitations

This study has several apparent strengths. The use of specific OSH questions from practice increases the ecological validity of our study. The stratified assignment of questions, the blinding of raters regarding the group to which the answers belonged, and the quasi-experimental design improved the quality of the study. Our study also has methodological limitations. The selection of workers who were planning to take part in an OSH supervisor course limits the generalizability to all workers. In addition, workers younger than 25 years were not represented in either group. As we believe that our sample may have been more motivated because of their proven OSH interest, the low rate of correct answers in the control group may be an overestimation.

Furthermore, although selecting OSH questions from ArboAntwoord may have increased the ecological validity, it also introduced a limitation. Participants probably did not personally relate to these specific OSH questions, and this might have caused participants to be less motivated to find an answer. This effect could be more apparent in the control group, who had to find an answer to these questions on their own. An alternative study design, letting participants bring in their own OSH question, may increase workers' understanding and commitment with the question, and as a consequence the efforts spent answering it. A disadvantage of this design is the potentially poor comparability of the outcomes between the two groups. The selection and composition of the questions might constitute another limitation. Our distinction between simple and complex questions can be questioned, as this was based on our personal estimation of whether an answer needed the combination and interpretation of information (complex). In retrospect, almost all our questions may be regarded as fairly complex, which might have caused the rate of correct answers to be lower than in daily practice. Furthermore, in view of classic evidence-based practice methods, at least several of the selected OSH questions seem poorly formulated. The accurate formulation of a clinical question is often mentioned as one of the most important skills required for evidence-based practice [9,10]. Several of our questions comprise more than one issue, do not address a specific target group, do not define a clear outcome, or do not take into account important contextual factors. Consequently, a poorly formulated question is more difficult to answer correctly, especially for participants in the control group, who had to interpret the question themselves. Moreover, this might also have influenced what we defined as a correct model answer, especially because for several questions the highest available level of evidence was the experts' (reviewers') opinion. In these cases, expert reviewers other than those involved in this study might have had a different opinion on what constitutes a correct evidence-based answer. In this line, the design and use of our model answers require consideration. For example, in this study we decided to dichotomize our ordinal data because of a skewed distribution. This may have caused some loss of information. Future studies may consider using continuous correctness scores. Lastly, our way of data collection may constitute a limitation. Participants had to answer the questions and complete the paper logs at home or at work. It might have been better to let participants search for information in the lab, where we could have videotaped them and assessed computer log files. However, the advantage of a field test is that it represents the real-life situation better than a laboratory experiment and that it is possible to study search strategies other than online ones.

Conclusions and Implications

Workers are often unable to find correct evidence-based answers to OSH questions when using common information sources, generally informational websites. The limited experience of workers with finding high-quality information seems to play an important role in finding correct answers; workers seem to be unable to judge the credibility, completeness, and applicability of the OSH information they find. Future research should explore workers' OSH information-seeking skills, their appraisal of information quality, and their ability to apply the obtained information to solve their question.

In addition to common information sources, other strategies and sources are required to assist workers in answering their OSH questions and to overcome difficulties in finding high-quality information. Expert advice provided through an online expert network (ArboAntwoord) can increase the rate of correct answers substantially, especially when focusing on the correct main conclusions. This purpose might also be facilitated by educational strategies such as short custom-made evidence-based practice courses for workers and managers or their representatives, or decision-support tools, or by providing accreditation to high-quality information. Future research could further establish the effectiveness of these new strategies.

Lastly, the identified difficulties with finding, appraising, and applying health-related information is not unique to workers. It is also relevant to other non-health professionals seeking health information, such as people in the general population or patients. Our findings on the potential value of online expert networks and expert facilities in general seem also applicable to other groups of people seeking answers to their health questions, albeit dependent on the quality of the knowledge infrastructures built around specific health topics (eg, asthma, cancer, or schizophrenia). Future research may focus on the impact of similar expert facilities in other health-related fields.

Acknowledgments

We thank the Foundation Institute GAK (Hilversum, The Netherlands) for funding this study. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors' work on this article was independent of the funder.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The 16 occupational safety and health (OSH) questions and their standardized correct model answers.

[PDF File (Adobe PDF File), 94KB - jmir_v14i1e9_app1.pdf]

References

1. Canadian Centre for Occupational Health and Safety. 2008. Report of the Council: April 1, 2007 to March 31, 2008 URL: http://www.ccohs.ca/ccohs/reports/annualReport0708.pdf [accessed 2011-06-21] [WebCite Cache ID 5zbc5tIfD]
2. Hoekstra P, van der Laan N. Inventarisatie naar informatiebehoeften over arbeidsomstandigheden [Inventory of Informational Needs in Occupational Health and Safety]. Report C6403. Amsterdam: TNS-NIPO; 2008.
3. Lang KH, Deilmann M, Nover H. Institut ASER e.V., Bergischen Universität Wuppertal, Germany. 2007. Zusammenfassung und Fortschreibung der Ergebnisse zum Pilotprojekt REACH-Net [Summary and Follow-up of the Results of the Pilot Project REACH-Net]. Forschungsbericht Nr 17 URL: http://www.institut-aser.de/pdf_files/literatur/REACH-Net-Zusammenfassung-Fortschreibung.pdf [accessed 2011-06-21] [WebCite Cache ID 5zbcNH3Ux]
4. Scott Porter Research and Marketing. NHS Health Scotland. 2004 May. Safe and Healthy Working: The Occupational Health Service for Small and Medium Enterprises (SMEs) URL: http://www.healthscotland.com/uploads/documents/RE04920032004Final.pdf [accessed 2011-06-21] [WebCite Cache ID 5zbcg8uw3]
5. Dryson E. Occupational health needs in small industry in New Zealand: preferred sources of information. Occup Med (Lond) 1993 Nov;43(4):176-179. [Medline: 8241474]
6. Hugenholtz NI, Schreinemakers JF, A-Tjak MA, van Dijk FJ. Knowledge infrastructure needed for occupational health. Ind Health 2007 Jan;45(1):13-18 [FREE Full text] [Medline: 17284868]
7. van Dijk FJ, Verbeek JH, Hoving JL, Hulshof CT. A knowledge infrastructure for occupational safety and health. J Occup Environ Med 2010 Dec;52(12):1262-1268. [doi: 10.1097/JOM.0b013e318202f2c5] [Medline: 21750475]
8. Shepperd S, Charnock D, Gann B. Helping patients access high quality health information. BMJ 1999 Sep 18;319(7212):764-766 [FREE Full text] [Medline: 10488009]
9. Offringa M, Assendelft WJJ. Inleiding in evidence-based medicine: klinisch handelen gebaseerd op bewijsmateriaal. 2nd edition. Houten: Bohn Stafleu Van Loghum; 2003.
10. Straus SE, Richardson WS, Glasziou P, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM. Edinburgh: Elsevier/Churchill Livingstone; 2005.
11. Anton B, Nelson R. Literacy, consumer informatics, and health care outcomes: Interrelations and implications. Stud Health Technol Inform 2006;122:49-53. [Medline: 17102216]
12. Schardt C. Health information literacy meets evidence-based practice. J Med Libr Assoc 2011 Jan;99(1):1-2. [doi: 10.3163/1536-5050.99.1.001] [Medline: 21243047]
13. World Health Organization. 1998. Health Promotion Glossary URL: http://www.who.int/hpr/NPH/docs/hp_glossary_en.pdf [accessed 2011-06-21] [WebCite Cache ID 5zbd1ITdj]
14. 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]
15. Haigh CA. Wikipedia as an evidence source for nursing and healthcare students. Nurse Educ Today 2011 Feb;31(2):135-139. [doi: 10.1016/j.nedt.2010.05.004] [Medline: 20646799]
16. Leithner A, Maurer-Ertl W, Glehr M, Friesenbichler J, Leithner K, Windhager R. Wikipedia and osteosarcoma: a trustworthy patients' information? J Am Med Inform Assoc 2010;17(4):373-374 [FREE Full text] [doi: 10.1136/jamia.2010.004507] [Medline: 20595302]
17. Tang H, Ng JH. Googling for a diagnosis--use of Google as a diagnostic aid: internet based study. BMJ 2006 Dec 2;333(7579):1143-1145 [FREE Full text] [doi: 10.1136/bmj.39003.640567.AE] [Medline: 17098763]
18. Hugenholtz NI, Schaafsma FG, Nieuwenhuijsen K, van Dijk FJ. Effect of an EBM course in combination with case method learning sessions: an RCT on professional performance, job satisfaction, and self-efficacy of occupational physicians. Int Arch Occup Environ Health 2008 Oct;82(1):107-115. [doi: 10.1007/s00420-008-0315-3] [Medline: 18386046]
19. Schaafsma F, Hugenholtz N, de Boer A, Smits P, Hulshof C, van Dijk F. Enhancing evidence-based advice of occupational health physicians. Scand J Work Environ Health 2007 Oct;33(5):368-378 [FREE Full text] [Medline: 17973063]

20. Burg MA, Zebrack B, Walsh K, Maramaldi P, Lim JW, Smolinski KM, et al. Barriers to accessing quality health care for cancer patients: a survey of members of the association of oncology social work. Soc Work Health Care 2010;49(1):38-52. [doi: 10.1080/00981380903018470] [Medline: 20077318]
21. Himmel W, Meyer J, Kochen MM, Michelmann HW. Information needs and visitors' experience of an Internet expert forum on infertility. J Med Internet Res 2005;7(2):e20 [FREE Full text] [doi: 10.2196/jmir.7.2.e20] [Medline: 15998611]
22. Iske P, Boersma W. Connected brains: question and answer systems for knowledge sharing: concepts, implementation and return on investement. J Knowl Manag 2005;9(1):126-145. [doi: 10.1108/13673270510583108]
23. Rhebergen MD, Hulshof CT, Lenderink AF, van Dijk FJ. An online network tool for quality information to answer questions about occupational safety and health: usability and applicability. BMC Med Inform Decis Mak 2010;10:63 [FREE Full text] [doi: 10.1186/1472-6947-10-63] [Medline: 20969777]
24. Coronel Institute, Dutch Center for Occupational Diseases, Academic Medical Center Amsterdam. 2011. ArboAntwoord URL: https://www.arboantwoord.com/default.aspx [accessed 2011-06-21] [WebCite Cache ID 5zbbhOlTU]
25. Thomas JR, Nelson JK. Research Methods in Physical Activity. Champaign, IL: Human Kinetics; 1996.
26. Hugenholtz NI, Nieuwenhuijsen K, Sluiter JK, van Dijk FJ. Do knowledge infrastructure facilities support evidence-based practice in occupational health? An exploratory study across countries among occupational physicians enrolled on Evidence-Based Medicine courses. BMC Health Serv Res 2009;9:18 [FREE Full text] [doi: 10.1186/1472-6963-9-18] [Medline: 19183450]
27. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
28. Hansen DL, Derry HA, Resnick PJ, Richardson CR. Adolescents searching for health information on the Internet: an observational study. J Med Internet Res 2003 Oct 17;5(4):e25 [FREE Full text] [doi: 10.2196/jmir.5.4.e25] [Medline: 14713653]
29. Peterson G, Aslani P, Williams KA. How do consumers search for and appraise information on medicines on the Internet? A qualitative study using focus groups. J Med Internet Res 2003 Dec 19;5(4):e33 [FREE Full text] [doi: 10.2196/jmir.5.4.e33] [Medline: 14713661]
30. Ilic D, Bessell TL, Silagy CA, Green S. Specialized medical search-engines are no better than general search-engines in sourcing consumer information about androgen deficiency. Hum Reprod 2003 Mar;18(3):557-561 [FREE Full text] [Medline: 12615824]
31. Minzer-Conzetti K, Garzon MC, Haggstrom AN, Horii KA, Mancini AJ, Morel KD, et al. Information about infantile hemangiomas on the Internet: how accurate is it? J Am Acad Dermatol 2007 Dec;57(6):998-1004. [doi: 10.1016/j.jaad.2007.06.038] [Medline: 17689833]
32. Whelan AM, Jurgens TM, Bowles SK, Doyle H. Efficacy of natural health products in treating osteoporosis: what is the quality of internet patient advice? Ann Pharmacother 2009 May;43(5):899-907. [doi: 10.1345/aph.1L688] [Medline: 19417114]
33. Thiele RH, Poiro 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]
34. Kingsley K, Galbraith GM, Herring M, Stowers E, Stewart T, Kingsley KV. Why not just Google it? An assessment of information literacy skills in a biomedical science curriculum. BMC Med Educ 2011;11:17 [FREE Full text] [doi: 10.1186/1472-6920-11-17] [Medline: 21518448]
35. Kortum P, Edwards C, Richards-Kortum R. The impact of inaccurate Internet health information in a secondary school learning environment. J Med Internet Res 2008;10(2):e17 [FREE Full text] [doi: 10.2196/jmir.986] [Medline: 18653441]
36. Schaafsma F, Verbeek J, Hulshof C, van Dijk F. Caution required when relying on a colleague's advice; a comparison between professional advice and evidence from the literature. BMC Health Serv Res 2005;5:59 [FREE Full text] [doi: 10.1186/1472-6963-5-59] [Medline: 16131405]
37. Harper FM, Raban D, Rafaeli S, Konstan JA. Predictors of answer quality in online Q&A sites. In: Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. 2008 Presented at: 26th SIGCHI Annual Conference on Human Factors in Computing Systems (CHI); 2008; Florence, Italy. [doi: 10.1145/1357054.1357191]
38. de Saintonge DM, Dunn DM. Gender and achievement in clinical medical students: a path analysis. Med Educ 2001 Nov;35(11):1024-1033. [Medline: 11703638]
39. Smits PB, Verbeek JH, Nauta MC, Ten Cate TJ, Metz JC, van Dijk FJ. Factors predictive of successful learning in postgraduate medical education. Med Educ 2004 Jul;38(7):758-766. [doi: 10.1111/j.1365-2929.2004.01846.x] [Medline: 15200400]

Abbreviations

CI: confidence interval
OSH: occupational safety and health

Edited by G Eysenbach; submitted 12.07.11; peer-reviewed by P Kortum; comments to author 28.09.11; revised version received 10.10.11; accepted 28.10.11; published 02.02.12.

Please cite as:
Rhebergen MDF, Lenderink AF, van Dijk FJH, Hulshof CTJ
Comparing the Use of an Online Expert Health Network against Common Information Sources to Answer Health Questions
J Med Internet Res 2012;14(1):e9
URL: http://www.jmir.org/2012/1/e9/
doi:10.2196/jmir.1886
PMID:22356848

Original Paper

Active-Q: Validation of the Web-Based Physical Activity Questionnaire Using Doubly Labeled Water

Stephanie Erika Bonn[1], MSc; Ylva Trolle Lagerros[2], MD, PhD; Sara Elisabeth Christensen[1], MSc; Elisabeth Möller[1], MSc; Antony Wright[3], BSC(Hons); Arvid Sjölander[1], PhD; Katarina Bälter[1], PhD

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[2]Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden

[3]MRC, Human Nutrition Research, Cambridge, United Kingdom

Corresponding Author:
Stephanie Erika Bonn, MSc
Department of Medical Epidemiology and Biostatistics
Karolinska Institutet
Nobels väg 12a
Stockholm, SE-171 77
Sweden
Phone: 46 8 524 822 98
Fax: 46 8 31 49 75
Email: stephanie.bonn@ki.se

Abstract

Background: Increased use of the Internet provides new opportunities for collecting data in large studies. The aim of our new Web-based questionnaire, Active-Q, is to assess total physical activity and inactivity in adults. Active-Q assesses habitual activity during the past year via questions in four different domains: (1) daily occupation, (2) transportation to and from daily occupation, (3) leisure time activities, and (4) sporting activities.

Objective: The objective of our study is to validate Active-Q's energy expenditure estimates using the doubly labeled water (DLW) method, and to assess the reproducibility of Active-Q by comparing the results of the questionnaire completed by the same group on two occasions.

Methods: The validity and reproducibility of Active-Q were assessed in a group of 37 individuals, aged 20 to 65 years. Active-Q was distributed via email to the participants. The total energy expenditure of the participants was assessed using DLW for 11 consecutive days.

Results: The median time to complete Active-Q was 6.1 minutes. The majority of participants (27/37, 73%) reported that the questionnaire was "easy" or "very easy" to answer. On average, Active-Q overestimated the total daily energy expenditure by 440 kJ compared with the DLW. The Spearman correlation between the two methods was r = 0.52 (P < .001). The intraclass correlation coefficient for total energy expenditure between the results of Active-Q completed on two occasions was 0.83 (95% CI 0.73-0.93).

Conclusions: Active-Q is a valid and reproducible method of assessing total energy expenditure. It is also a user-friendly method and suitable for Web-based data collection in large epidemiological studies.

(J Med Internet Res 2012;14(1):e29) doi:10.2196/jmir.1974

KEYWORDS

activity assessment; epidemiology; Internet; total energy expenditure

Introduction

Increased use of the Internet over the past decade provides new opportunities to use Web-based questionnaires in epidemiological studies. Compared to traditional paper-based questionnaires, Web-based alternatives may simplify data collection and improve the quality of data [1-3]. The need to design and validate questionnaires specifically for use on the Web has, therefore, increased in recent years. Many Web-based questionnaires used today are directly transferred from printed questionnaires rather than being originally developed for the medium in which they are used. Thus, few take advantage of

the potential interactive features offered by the Web. For example, the burden on the respondent can be decreased by using follow-up questions, and errors can be minimized by implementing automatic controls for missing, inconsistent, or anomalous answers [4].

At present, there are few physical activity questionnaires available specifically developed for the Web. To the best of our knowledge, only one Web-based questionnaire (that assessed lifetime physical activity) has been validated and showed acceptable results [5]. Therefore, we developed a series of interactive questionnaires that assess total physical activity and inactivity in different age groups. The Active-Q questionnaire assesses physical activity and inactivity during the past year in subjects over 18 years.

The primary aim of this study was to test the validity of Active-Q against the doubly labeled water (DLW) method, the criterion standard for measuring energy expenditure [6]. Furthermore, the reproducibility of Active-Q was assessed by comparing the results obtained from the questionnaire on two separate occasions.

Methods

Study Participants

Study participants of both sexes, aged 20 to 65 years, were recruited in April of 2009 through public advertisements (including advertisements on the campuses of three universities) around Stockholm, Sweden. Participants were required to have an email address and access to the Internet. Exclusion criteria were any form of weight alteration diet, pregnancy, or having given birth during the ten months prior to the start of the study. Participants were provided with written and verbal information about the study. All participants gave their written informed consent prior to entering the study.

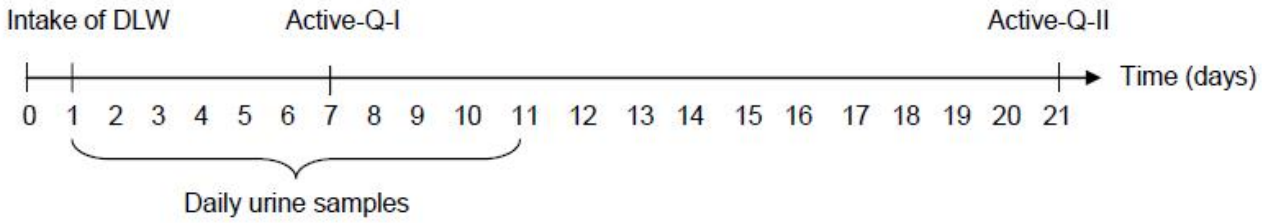
In total, 40 individuals were recruited. Data from three participants were excluded from the analysis because of illness during data collection, incomplete data from the first Active-Q questionnaire (ie, < 1 hour of leisure time activities per day reported), and unreliable DLW data, respectively. After exclusions, data from 37 participants remained for analysis.

Study Design

Figure 1 illustrates the design of the study procedure. Participants filled out the Active-Q questionnaire on two occasions, separated by two weeks, and their energy expenditure was measured for 11 consecutive days using the DLW method [7]. On day 1, a baseline urine sample was collected before the participants drank a dose of DLW. Daily urine samples were collected for the next 10 consecutive days. Participants received the first Active-Q questionnaire (Active-Q-I) 7 days after the start of the study; they received the second questionnaire (Active-Q-II) two weeks later. Questionnaires were distributed via email. Individual usernames and passwords served as unique identifiers in the Web-questionnaire program. Email reminders were sent to participants who had not responded to the questionnaire within two days of it being sent out. After responding to Active-Q-I, participants were asked to evaluate its user-friendliness and self-report their weight, height, level of education, and tobacco use.

The study was approved by the Research Ethics Committee at the Karolinska Institutet, Stockholm, Sweden.

Figure 1. Timeline of study showing days that participants took doubly labeled water (DLW), provided urine samples, took the first questionnaire (Active-Q-I), and the second questionnaire (Active-Q-II).



Active-Q

Active-Q is a Web-based questionnaire designed to assess physical activity and inactivity in adults older than 18 years. Respondents are asked to report their habitual activity during the past year. Active-Q covers four different domains: (1) daily occupation, (2) transportation to and from daily occupation, (3) leisure time activities, and (4) regular sporting activities (see Multimedia Appendix 1). The leisure time domain includes activities such as housework, watching television, and using the computer. Within each domain, activities with similar intensity levels are grouped together as appropriate. Inactivity is assessed by questions about sedentary behavior during daily occupation, and by questions regarding time spent watching television and using a computer during leisure time. The initial questions within the means of transportation to and from daily occupation, leisure time activities, and sporting activities domains are screening questions that list all the activities included in the domain. The participant selects those activities within this list that he or she practices. Follow-up questions about frequency and duration pertain only to those activities selected in the screening question. With the exception of one question (the total number of hours spent in daily occupation), all questions have predefined answers regarding frequency and duration. Respondents report the frequency and duration of leisure time activities they perform at least once per week and sport activities they perform on a regular basis, thus limiting the number of questions for each respondent. Figure 2 shows screenshots of an initial screening question in Active-Q and Figure 3 shows a screenshot of a follow-up question for a selected activity regarding frequency and duration.

In total, Active-Q includes 35 questions with all activities linked to a corresponding Metabolic Equivalent Task (MET) value [8, 9]. Table 1 lists the activities included in Active-Q and their corresponding MET values.

Energy expenditure (EE) was estimated based on the assumption that 1 MET equals 1 kcal $\cdot$ kg^{-1} $\cdot$ h^{-1} [8]. To calculate the energy expenditure based on the results from Active-Q, MET values assigned to each activity were multiplied by the participants' weight (kg), the average daily duration of the activity (h/d), and a conversion factor of 4.184, to transform the values from kcal to kJ:

$$EE_{activity}\ (kJ/day) = MET_{activity} \cdot Weight\ (kg) \cdot Duration_{activity}\ (h/d) \cdot 4.184$$

Total energy expenditure was expressed in terms of the crude total energy expenditure and the total energy expenditure per 24 hours. The crude total energy expenditure was obtained by summing the contributions of energy from each activity in Active-Q. The total energy expenditure over a 24-hour period was calculated by adding eight hours of sleep to the crude results from Active-Q. If the resulting time differed from 24 hours, time was added or subtracted to obtain the adjusted total energy expenditure per 24 hours. A MET value of 2.0 was assumed for the time added and subtracted [8, 10].

Figure 2. Screenshot of Active-Q screening question regarding mode of transportation to and from daily activities.

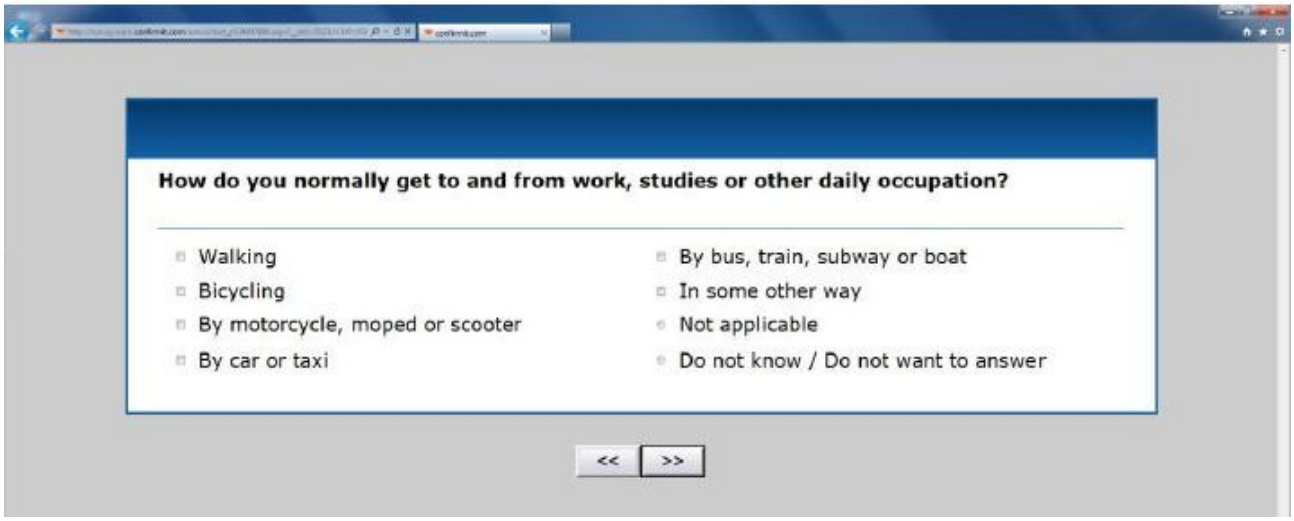


Figure 3. Screenshot of follow-up question for activities selected in screening question regarding transportation showing an example of possible answers.

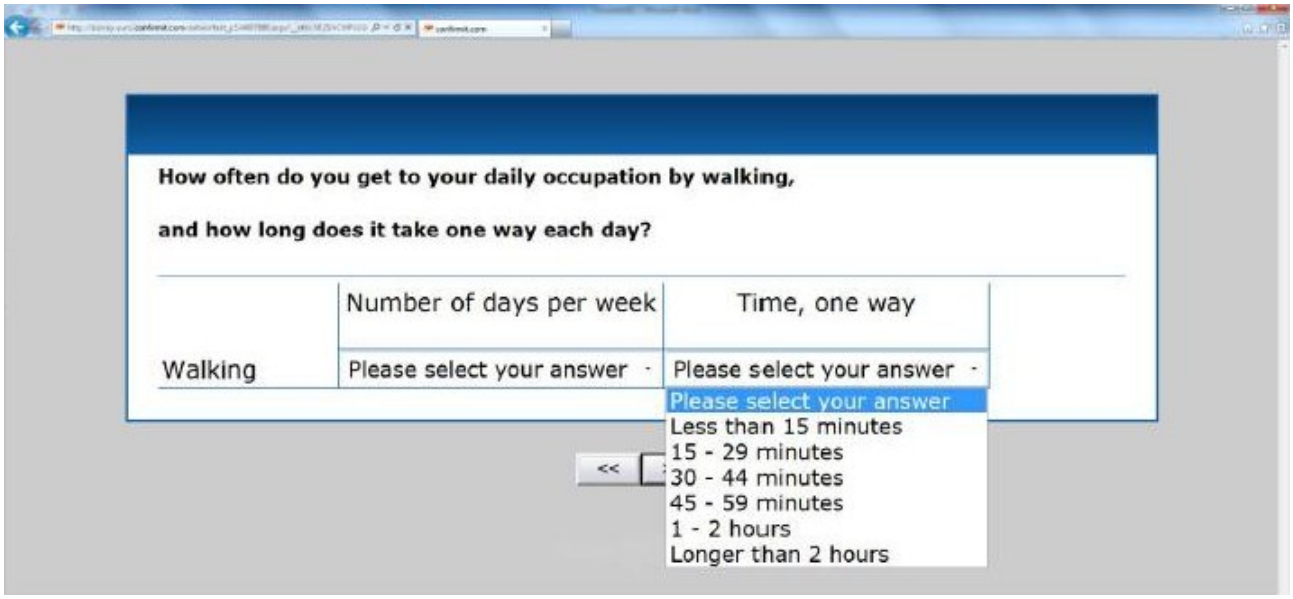


Table 1. Questions included in Active-Q[a] and corresponding MET values.

Activity category		MET value
Daily occupation [b]		
	Mostly sitting	1.5
	A combination of sitting and standing up	2.3
	Mostly standing up	3.0
	Some physical activity	4.5
	Heavy manual labor	6.0
Transportation		
	Walking	4.0
	Bicycling	4.0
	By motorcycle or scooter	2.5
	By car or taxi	1.0
	By bus, train, subway, or boat	1.0
Leisure time activities		
	Watching TV/DVDs	1.0
	Using the computer	1.0
	Sitting listening to music, sewing, etc	1.0
	Playing a musical instrument or active computer games	2.0
	Doing household chores	3.0
	Shopping or other errands	2.3
	Dancing	3.0
	Walking	3.4
	Bicycling	8.0
Regular sporting activities		
	Aerobics	6.5
	Weight lifting	6.0
	Jogging or running	8.0
	Athletics	6.0
	Spinning	8.5
	Swimming	6.0
	Soccer, basketball, volleyball, or floorball (floor hockey)	6.0
	Golf	4.5
	Dance class	4.5
	Horseback riding	4.0
	Ice skating, ice hockey, or bandy	7.0
	Skiing (downhill or cross country)	7.0
	Martial arts	10.0
	Boxing or wrestling	6.0
	Tennis, badminton, or squash	7.0
	Table tennis	4.0
	Rowing, canoeing, surfing, or sailing	3.0
	Motor sports	4.0

Activity category		MET value
	Rock climbing	8.0
	Other	2.5

[a] All domain activities (ie, daily occupation, transportation, leisure time, and sporting activities), including frequency and duration, were assessed via an initial screening questionnaire.

[b] Participants ranked their overall effort in this category on a scale from 1 to 5.

Doubly Labeled Water

Doubly labeled water was used as the criterion measure of total energy expenditure [7] for 11 consecutive days. The logistics of the study prevented the preparation of individually tailored DLW doses. Instead, two standard doses were prepared adopting a strategy similar to that of Trabulsi et al [11]. A bulk dose of DLW was made by adding 44 g of ^2H_2O (99.98% sterility tested, CK Gas Products Ltd, Hampshire, UK) to 1 L 10% normalized $H_2{}^{18}O$ (SerCon Ltd, Cheshire, UK). Participants gave a baseline urine sample before drinking either 108 g (participants < 75 kg) or 141 g (participants ≥ 75 kg) of the bulk dose depending on their self-reported weight at the beginning of the study. In addition, participants collected 5 ml daily urine samples (excluding the first morning void) on each of the following 10 days. Participants recorded the date and time of each sample collection and kept the samples refrigerated until they were returned to the research team. The research team sent the samples to MRC Human Nutrition Research in Cambridge, UK, for analysis.

The principles for the slightly modified analyses of the isotopic enrichment in the samples have been described in detail in previous studies [12, 13]. Briefly, for ^2H analysis, 0.4 ml samples of undistilled urine were equilibrated with hydrogen gas using platinum catalyst rods to promote the rapid exchange between ^2H and ^1H. The 3.5-ml sample vials were first flush-filled with hydrogen, producing approximately 3 ml of gas at 1 bar atmosphere, and then equilibrated at 22.0 +0.1°C for 6 hours. Cryogenically dried hydrogen gas from the headspaces of each sample was analyzed using isotope ratio mass spectrometry (Sira10, VG Isogas, Winsford, Cheshire, UK).

Measurements of the $^{18}O/^{16}O$ ratios were made using an AP2003 continuous-flow isotope ratio mass spectrometer (Analytical Precision Ltd, Northwich, Cheshire, UK). Urine samples of 0.5 ml were placed in 10 ml Vacutainers (Labco Ltd, High Wycombe, UK), flush-filled with 5% CO_2 in nitrogen and then equilibrated on blood tube rotators overnight at room temperature before analysis. In all cases, analytical standards prepared in-house and traceable to the international standards, Vienna Standard Mean Ocean Water (VSMOW) and Standard Light Arctic Precipitation (SLAP), were included in each batch of samples analyzed.

Total energy expenditure was calculated as described by Schoeller et al [14] from the slopes and intercepts of the isotope disappearance curves using urine samples collected on days 1-3 and 8-10, respectively. The value of the respiratory quotient was assumed to be 0.85.

Statistical Methods

Descriptive statistics were used to present the characteristics of the participants. Results are reported as mean values and standard deviations (SD), with the exception of time to respond to the questionnaire, which is presented as the median response time, and results of user-friendliness, which is presented in absolute numbers. For categorical variables or variables with a skewed distribution, Fisher's exact test was performed to assess potential differences between men and women, and participants < 30 and ≥ 30 years of age. For those continuous variables typically normally distributed (eg, height, weight, and BMI), *t* tests were used to assess potential differences. The level of significance was set at $P < .05$.

The degree of association between the total energy expenditure obtained from Active-Q-I, both crude and adjusted to reflect a 24-hour day, and the DLW method, was assessed using Spearman correlation coefficients. Because Spearman correlation coefficients do not detect systematic differences between the methods, we used the Bland-Altman technique [15] to determine the absolute agreement. The difference in energy expenditure assessed from Active-Q-I (adjusted to reflect a 24-hour period) and DLW was plotted on the y-axis, and the mean values of the two assessments on the x-axis. The limits of agreement, equal to ±2 SD of the mean difference, provide a measure of the variation. The reproducibility of Active-Q was assessed using intraclass correlation coefficients [16]. The ANOVA estimator of intraclass correlation coefficients was computed. All analyses were performed using STATA 11.1 (STATA Corporation, College Station, TX).

Results

Table 2 displays the baseline characteristics of the participants included in analysis. Subjects were predominantly female (> 80%) and under the age of 40 (> 70%). The self-reported height and weight resulted in a mean body mass index (BMI) of 23.0 (±3.8) kg/m^2. Men were taller ($P < .001$) and weighed more ($P < .001$) than women. Subjects ≥ 30 years of age, weighed more ($P = .024$) and had a higher BMI ($P = .007$) compared with participants < 30 years of age. No other statistically significant differences were found between sex or age groups.

The median time required to fill out the Active-Q questionnaire on the first occasion was 6 minutes and 6 seconds. On average, the activities reported corresponded to 11 hours and 24 minutes of a typical day (excluding time spent sleeping). In the evaluation of the user-friendliness of the questionnaire, the majority of respondents (27/37, 73%) graded the questionnaire as "easy" or "very easy" to answer. Respondents were also asked to give the questionnaire an overall grade on a scale from 1 (worst) to 5 (best). The majority of respondents (21/37, 57%)

graded the questionnaire a 4. An equal number of participants (7/37, 19%), graded the questionnaire a 3 or a 5. The remainder of participants (2/37, 5%) gave the questionnaire a 2. No respondents gave the questionnaire a 1, the worst possible grade. The mean overall grade for Active-Q-I was 3.9 ±0.8. Participants ≥ 30 years graded the questionnaire higher than younger participants (P = .006). No other statistically significant age-related or sex-related differences were seen in the results.

The mean total daily energy expenditure measured with DLW was 11,229 kJ (SD 2256), while the mean energy expenditure from Active-Q-I and Active-Q-II, adjusted to reflect a 24-hour period, was 11,667 kJ (SD 3212) and 11,529 kJ (SD 2758), respectively. The mean crude energy expenditure assessed with Active-Q-I was 7008 kJ (SD 3854). The crude energy expenditure for each domain of Active-Q-I was 2971 kJ (SD 1736) for daily occupation, 434 kJ (SD 388) for transportation, 2243 kJ (SD 1550) for leisure time activities, and 1360 kJ (SD 3044) for regular sporting activities. The mean crude energy expenditure from Active-Q-II was 6439 kJ (SD 2614). The crude energy expenditure for each domain of Active-Q-II was 3005 kJ (SD 1537) for daily occupation, 365 kJ (SD 296) for transportation, 2254 kJ (SD 1880) for leisure time activities, and 815 kJ (SD 761) for regular sporting activities.

The Spearman correlation coefficient between the crude total daily energy expenditure assessed with Active-Q-I and DLW was r = 0.42 (P = .01). After adjusting Active-Q-I for a 24-hour day, the Spearman correlation coefficient increased to r = 0.52 (P < .001).

Figure 4 shows a Bland-Altman plot illustrating the differences in the total energy expenditure during 24 hours assessed with Active-Q-I and DLW. No clear trend (proportional error) can be seen in the plot. Most data points fall within the limits of agreement (±2 SD), although it should be noted that the limits of agreement were wide. The mean difference between the methods of assessment was 440 kJ, with Active-Q overestimating the total daily energy expenditure compared with DLW. According to the Bland-Altman plot, there were two outliers for which Active-Q noticeably overestimated total energy expenditure. One of the outliers reported extreme amounts of sporting activities in Active-Q, while the other had a BMI above 35 kg/m^2.

The intraclass correlation coefficient for the crude total energy expenditure assessed from the questionnaire on two occasions was 0.66 (95% CI 0.47-0.84). After adjustments to a 24-hour day, the value increased to 0.83 (95% CI 0.73-0.93).

Figure 4. Bland-Altman plot illustrating the difference in total daily energy expenditure (TEE) between Active-Q-I (adjusted to reflect a 24-hour day) and the DLW method. The absolute difference in TEE between Active-Q-I and DLW is plotted on the y-axis and the mean of the two assessments on the x-axis. Each data point represents one participant (n= 37).

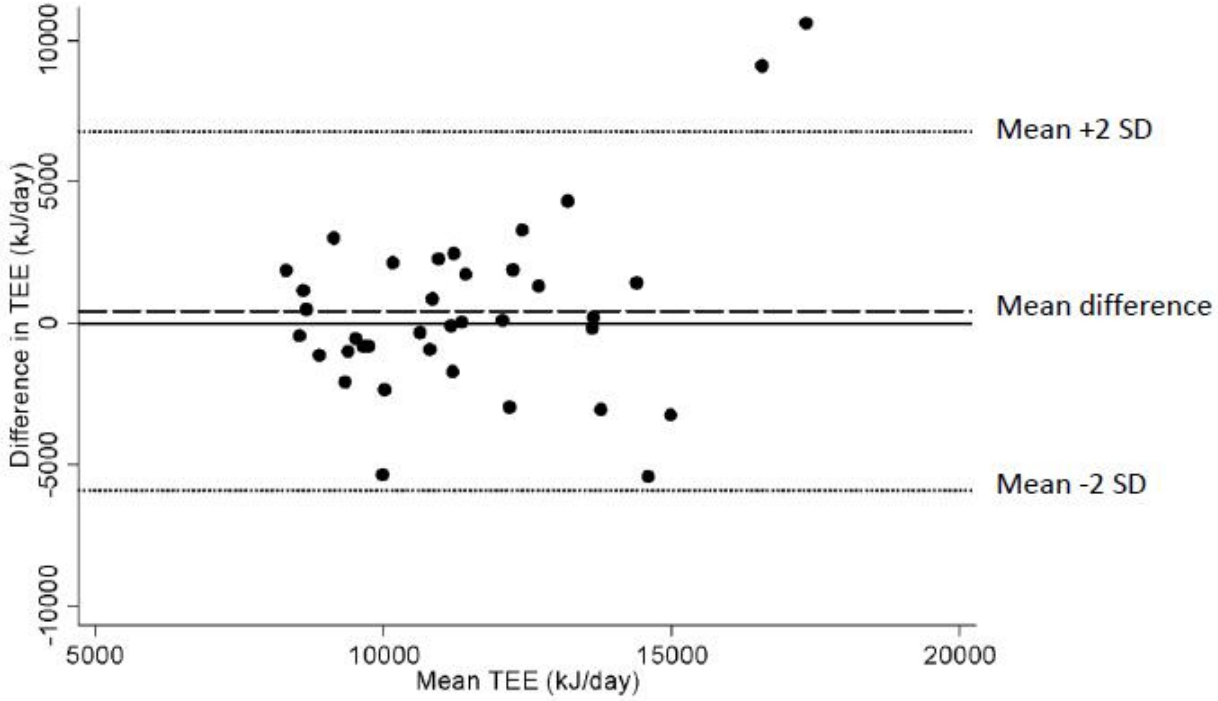


Table 2. Baseline characteristics of the study population (n = 37).

	n	(%)
Sex		
Female	30	(81)
Male	7	(19)
Age (y)		
< 30	22	(59)
30-39	5	(14)
40-49	5	(14)
50-59	4	(11)
> 60	1	(3)
BMI (kg/m^2)		
< 20	3	(8)
20-25	28	(76)
> 25	6	(16)
Education (y)		
9-12	7	(19)
> 12	30	(81)
Smoking status[a]		
Current	2	(5)
Previous	8	(22)
Never	25	(68)
Snuff use[a]		
Current	4	(11)
Previous	6	(16)
Never	26	(70)
Watching television[b] **(h/d)**		
< 1	25	(68)
1-3	10	(27)
≥ 3	2	(5)
Computer use[b, c] **(h/d)**		
< 1	23	(62)
1-3	8	(22)
≥ 3	6	(16)
Regular sport activities[b] **(h/d)**		
0	5	(14)
< 1	26	(70)
1-3	5	(14)
≥ 3	1	(3)

[a] Percentages do not equal 100 because of missing data

[b] Reported in Active-Q-I

[c] During leisure time

Discussion

The results of this study demonstrate that Active-Q is a user-friendly questionnaire that provides valid and reproducible estimates of total energy expenditure when compared with objective DLW measurements, the criterion standard [6].

The validity and reproducibility of many paper-based physical activity questionnaires in use today are low [17], although some show reasonable validity [18-22]. The number of questionnaires validated against DLW is limited and, according to two reviews, few studies have shown Spearman correlations above 0.50 [23, 24]. However, Besson et al [20] recently reported a correlation of 0.67. Nevertheless, all questionnaires previously validated against DLW have, to the best of our knowledge, been paper-based and not adapted for the Web, making comparisons with Active-Q difficult.

Important factors in making a questionnaire user-friendly are the length and the level of details in a questionnaire, plus the way in which questions are structured and the order in which they are presented [25]. It is also important to adapt the questionnaire to the medium in which it is to be used for data collection. Web-based questionnaires offer many advantages compared with traditional paper-based questionnaires. The Web allows the use of skip and follow-up patterns, thereby increasing the user-friendliness by creating interactivity, which may reduce dropout [3, 26]. Technical features, such as checking the plausibility of answers, decrease the measurement error [27]. Computerized administration also means reduced costs for data collection and analysis, and facilitates the administration of large studies [1, 26]. Active-Q is interactive and takes advantage of follow-up questions with predefined answers, which minimizes the time required to answer the questions and decreases the risk of mistyping. Further, checks for missing answers are implemented in Active-Q, and respondents are prevented from proceeding to the next question in the questionnaire before answering the previous questions.

Today, more than 90% of the population in Sweden between 16 and 74 years of age have access to the Internet at home [28]. Therefore, we do not believe that concerns raised previously regarding high non-response rates and the introduction of selection bias will be a problem when collecting Web-based data in Sweden [26]. However, it is difficult to draw any conclusions regarding non-response rates and selection bias from this study, as a pre-requisite for participation was access to the Internet and an email address. A previous study comparing response rates to Web- and paper-based dietary questionnaires in Sweden reported that the response rate for paper questionnaires was approximately 15% higher than for Web questionnaires [3]. However, this study was performed in 2002 and access to, and the use of, the Internet have increased considerably in Sweden during the past decade [28].

Among Swedish individuals between 16 and 44 years of age, over 97% reported to be frequent users of the Internet compared to 90% for individuals between 45 and 54 years, and 70% for individuals between 55 and 74 years [28]. Although younger individuals are more frequent users of the Internet, the vast majority of the older individuals also use the Internet regularly. Therefore, we believe that a Web-based questionnaire is suitable for a wide age range. However, a larger validation study would clarify how applicable a Web-based questionnaire would be across different age groups.

Ideally, in validation studies, the reference method should reflect the same period of time as the new method being validated. However, habitual physical activity was assessed over a one-year period with Active-Q, while that measured with the DLW method reflects physical activity over 11 days. Repeated reference measures of DLW over a longer period of time would have been preferable, but this was neither practical nor possible in our study setting. Another drawback of this study is that we were only able to validate measures of total energy expenditure. Because the DLW method does not discriminate between different levels of intensity in activity, we were not able to validate Active-Q in this respect. Since different aspects of activity may affect health-related outcomes differently, the ability of Active-Q to accurately assess this needs to be validated in future studies using, for example, accelerometers.

When estimating the total daily energy expenditure with Active-Q, we assumed that all participants slept for eight hours [29]. The true variation in the duration of sleep among individuals may reduce the accuracy of results, leading to an underestimation of validity of Active-Q. In future studies, a question about the average duration of sleep will be added to Active-Q. On average, the participants' reported total activity time in Active-Q-I was in line with other studies on the Swedish population [30]. However, even after adjusting for sleep, most participants underestimated the time spent in daily activities in Active-Q, whereas a small proportion overestimated it. Therefore, results were adjusted by adding or subtracting under- or over-reported time to obtain a total of 24 hours in order to make the data comparable to the average total daily energy expenditure obtained from DLW measurements. Missing and overestimated time was adjusted by using a MET value of 2.0. Missing time was assumed to reflect a combination of common activities, such as self-care and preparing and eating meals, and overestimated time to reflect an average MET value of activities included in Active-Q.

The use of published MET values [9] to represent the energy cost for a specific activity constitutes a limitation in most physical activity questionnaires. The MET value for a specific activity assumes the same energy expenditure per kilogram of body weight for all individuals, regardless of variations in both mechanical and metabolic efficiency [8, 9], thereby leading to the risk of potential misclassification in assessments of energy expenditure. However, the measure of physical activity used in many epidemiological studies is that of MET hours, avoiding the issue of weight being a factor affecting the final outcome in terms of energy expenditure, which is obtained by multiplying MET hours with participants' weight. Hence, the exact energy expenditure of participants is less important than ranking individuals correctly.

While the point estimate of the intraclass correlation for the adjusted total energy expenditure assessed with Active-Q is high, the point estimate for the assessed crude total energy expenditure is lower, and the confidence intervals around both

estimates are rather wide. The wide confidence intervals, and the rather low point estimate obtained for the crude total energy expenditure, might be explained by our small sample size and sampling variability. To obtain more certain measures of reproducibility, we need to conduct a larger study to assess reproducibility of both assessments of total energy expenditure, as well as energy expenditure within each individual domain of Active-Q.

Our study population was recruited from the Stockholm area, including the campuses of three universities, and the participants were young and predominantly female. This may limit the generalizability of our results. The overall level of education among the participants was high, with more than 80% having an education longer than 12 years, compared with 36% in the general Swedish population aged 16-64 years [31]. Higher education has been associated with more accurate recall of physical activity [32]; therefore, our results may be biased in this direction. In addition, bias due to self-selection of participants may be present because recruitment was voluntary through public advertisements. Health-conscious individuals may be more prone to participate in a study on physical activity.

To the best of our knowledge, Active-Q is one of the first validated physical activity questionnaires specifically designed for Web-based use. Despite the limitations mentioned previously, we believe Active-Q to be a good assessment method in studies collecting Web-based data. Active-Q is currently in use in three large ongoing cohort studies in Sweden, aimed at recruiting hundreds of thousands of people [33].

An increasing number of studies are relying on the collection of data via the Web, and there is a need for valid Web-based methods of assessing physical activity. We have demonstrated that Active-Q is a valid method for estimating total energy expenditure. Furthermore, Active-Q is also a reproducible and user-friendly method. We conclude that Active-Q is a suitable method for collecting Web-based data on physical activity and inactivity in large epidemiological studies.

Acknowledgments

This work was supported by grants from the Swedish Cancer Foundation as well as grants to LifeGene from AFA Insurance, Torsten and Ragnar Söderbergs Foundation, Karolinska Institutet, Stockholm's County Council, and the Swedish Research Council.

The authors thank MRC Human Nutrition Research for their cooperation and performing the DLW analyses, in particular Marilena Leventi for her help in analyzing samples and Dr. Les Bluck for his help with the analysis and interpretation of the results.

Authors' Contributions

SEB, YTL, SEC, EM, and KB contributed significantly to the concept and design of the study, and SEB, SEC, EM, AW, and KB participated in the data acquisition. AW conducted doubly labeled water analyses while SEB conducted data analyses (in collaboration with AS who provided significant statistical expertise). SEB, YTL, AW, AS, and KB contributed to the interpretation of results. SEB together with YTL and KB prepared the manuscript, which was subsequently reviewed by SEC, EM, AW, and AS. All authors have approved the final manuscript.

Multimedia Appendix 1

The Active-Q Physical Activity Questionnniare.

[PDF File (Adobe PDF File), 300KB - jmir_v14i1e29_app1.pdf]

References

1. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
2. Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using web-based questionnaires in large population-based epidemiological studies. Eur J Epidemiol 2006;21(2):103-111. [doi: 10.1007/s10654-005-6030-4] [Medline: 16518678]
3. Bälter KA, Bälter O, Fondell E, Lagerros YT. Web-based and mailed questionnaires: a comparison of response rates and compliance. Epidemiology 2005 Jul;16(4):577-579. [Medline: 15951679]
4. 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]
5. De Vera MA, Ratzlaff C, Doerfling P, Kopec J. Reliability and validity of an internet-based questionnaire measuring lifetime physical activity. Am J Epidemiol 2010 Nov 15;172(10):1190-1198 [FREE Full text] [doi: 10.1093/aje/kwq273] [Medline: 20876666]
6. Ainslie P, Reilly T, Westerterp K. Estimating human energy expenditure: a review of techniques with particular reference to doubly labelled water. Sports Med 2003;33(9):683-698. [Medline: 12846591]

7. Davidsson L. Assessment of body composition and total energy expenditure in humans using stable isotope techniques. IAEA Human Health Series. Vienna: International Atomic Energy Agency; 2009. URL: http://www-pub.iaea.org/MTCD/publications/PDF/Pub1370_web.pdf [accessed 2011-10-25] [WebCite Cache ID 62hU5ssrg]
8. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR, Montoye HJ, Sallis JF, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc 1993 Jan;25(1):71-80. [Medline: 8292105]
9. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S498-S504. [Medline: 10993420]
10. Norman A, Bellocco R, Bergström A, Wolk A. Validity and reproducibility of self-reported total physical activity--differences by relative weight. Int J Obes Relat Metab Disord 2001 May;25(5):682-688 [FREE Full text] [doi: 10.1038/sj.ijo.0801597] [Medline: 11360151]
11. Trabulsi J, Troiano RP, Subar AF, Sharbaugh C, Kipnis V, Schatzkin A, et al. Precision of the doubly labeled water method in a large-scale application: evaluation of a streamlined-dosing protocol in the Observing Protein and Energy Nutrition (OPEN) study. Eur J Clin Nutr 2003 Nov;57(11):1370-1377. [doi: 10.1038/sj.ejcn.1601698] [Medline: 14576749]
12. Assah FK, Ekelund U, Brage S, Wright A, Mbanya JC, Wareham NJ. Accuracy and validity of a combined heart rate and motion sensor for the measurement of free-living physical activity energy expenditure in adults in Cameroon. Int J Epidemiol 2011 Feb;40(1):112-120 [FREE Full text] [doi: 10.1093/ije/dyq098] [Medline: 20529884]
13. Bluck LJC. Doubly labelled water for the measurement of total energy expenditure in man - progress and applications in the last decade. Nutrition Bulletin 2008;33:80-90. [doi: 10.1111/j.1467-3010.2008.00695.x]
14. Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P, Jéquier E. Energy expenditure by doubly labeled water: validation in humans and proposed calculation. Am J Physiol 1986 May;250(5 Pt 2):R823-R830. [Medline: 3085521]
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
16. Lee J, Koh D, Ong CN. Statistical evaluation of agreement between two methods for measuring a quantitative variable. Comput Biol Med 1989;19(1):61-70. [Medline: 2917462]
17. Westerterp KR. Assessment of physical activity: a critical appraisal. Eur J Appl Physiol 2009 Apr;105(6):823-828. [doi: 10.1007/s00421-009-1000-2] [Medline: 19205725]
18. Conway JM, Irwin ML, Ainsworth BE. Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires - a doubly labeled water validation. J Clin Epidemiol 2002 Apr;55(4):392-399. [Medline: 11927208]
19. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
20. Besson H, Brage S, Jakes RW, Ekelund U, Wareham NJ. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. Am J Clin Nutr 2010 Jan;91(1):106-114 [FREE Full text] [doi: 10.3945/ajcn.2009.28432] [Medline: 19889820]
21. Staten LK, Taren DL, Howell WH, Tobar M, Poehlman ET, Hill A, et al. Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. Med Sci Sports Exerc 2001 Nov;33(11):1959-1967. [Medline: 11689750]
22. Pettee Gabriel K, McClain JJ, Schmid KK, Storti KL, Ainsworth BE. Reliability and convergent validity of the past-week Modifiable Activity Questionnaire. Public Health Nutr 2010 Sep 15:1-8. [doi: 10.1017/S1368980010002612] [Medline: 20843404]
23. Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? Am J Clin Nutr 2008 Feb;87(2):279-291 [FREE Full text] [Medline: 18258615]
24. van Poppel MN, Chinapaw MJ, Mokkink LB, van Mechelen W, Terwee CB. Physical activity questionnaires for adults: a systematic review of measurement properties. Sports Med 2010 Jul 1;40(7):565-600. [doi: 10.2165/11531930-000000000-00000] [Medline: 20545381]
25. Jacobs DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc 1993 Jan;25(1):81-91. [Medline: 8423759]
26. Russell CW, Boggs DA, Palmer JR, Rosenberg L. Use of a web-based questionnaire in the Black Women's Health Study. Am J Epidemiol 2010 Dec 1;172(11):1286-1291 [FREE Full text] [doi: 10.1093/aje/kwq310] [Medline: 20937635]
27. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
28. Arrhenius C, Ewerdahl D. Statistics Sweden. 2011. Use of computers and the Internet by private persons in 2010 URL: http://www.scb.se/statistik/_publikationer/LE0108_2010A01_BR_IT01BR1101.pdf [accessed 2011-10-18] [WebCite Cache ID 62Wgb5Nr0]
29. Rydenstam K. Statistics Sweden. 2003. Living conditions, report no. 99. Time for everyday life: Women's and men's time use 1990/91 and 2000/01 URL: http://www.scb.se/statistik/LE/LE0103/2003M00/LE99SA0301.pdf [accessed 2011-10-18] [WebCite Cache ID 62Wh0f4el]
30. Lagerros YT, Bellocco R, Adami HO, Nyrén O. Measures of physical activity and their correlates: the Swedish National March Cohort. Eur J Epidemiol 2009;24(4):161-169. [doi: 10.1007/s10654-009-9327-x] [Medline: 19294519]

31. Sundh L. Statistics Sweden. 2010. Yearbook of Educational Statistics 2011 URL: http://www.scb.se/statistik/_publikationer/UF0524_2010A01_BR_00_UF0110TAB.pdf [accessed 2011-09-26] [WebCite Cache ID 61zB2vUXx]
32. Falkner KL, McCann SE, Trevisan M. Participant characteristics and quality of recall of physical activity in the distant past. Am J Epidemiol 2001 Nov 1;154(9):865-872 [FREE Full text] [Medline: 11682369]
33. Almqvist C, Adami HO, Franks PW, Groop L, Ingelsson E, Kere J, et al. LifeGene--a large prospective population-based study of global relevance. Eur J Epidemiol 2011 Jan;26(1):67-77. [doi: 10.1007/s10654-010-9521-x] [Medline: 21104112]

Abbreviations

BMI: body mass index
DLW: doubly labeled water
EE: energy expenditure
MET: metabolic equivalent task
SLAP: standard light arctic precipitation
VSMOW: Vienna standard mean ocean water

Edited by G Eysenbach; submitted 26.10.11; peer-reviewed by P Bergman, S Blair, M De Vera; comments to author 27.11.11; revised version received 20.12.11; accepted 02.01.12; published 15.02.12.

Please cite as:
Bonn SE, Trolle Lagerros Y, Christensen SE, Möller E, Wright A, Sjölander A, Bälter K
Active-Q: Validation of the Web-Based Physical Activity Questionnaire Using Doubly Labeled Water
J Med Internet Res 2012;14(1):e29
URL: http://www.jmir.org/2012/1/e29/
doi:10.2196/jmir.1974
PMID:22356755

Original Paper

HealthTrust: A Social Network Approach for Retrieving Online Health Videos

Luis Fernandez-Luque[1,2], MSc; Randi Karlsen[1,2], PhD; Genevieve B Melton[3], MA, MD

[1]Northern Research Institute, Tromsø, Norway

[2]Computer Science Department, University of Tromsø, Tromsø, Norway

[3]Institute for Health Informatics, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
Luis Fernandez-Luque, MSc
Northern Research Institute
Postboks 6434 Forskningsparken
Tromsø, 9294
Norway
Phone: 47 93421287
Fax: 47 77629401
Email: luis.luque@norut.no

Abstract

Background: Social media are becoming mainstream in the health domain. Despite the large volume of accurate and trustworthy health information available on social media platforms, finding good-quality health information can be difficult. Misleading health information can often be popular (eg, antivaccination videos) and therefore highly rated by general search engines. We believe that community wisdom about the quality of health information can be harnessed to help create tools for retrieving good-quality social media content.

Objectives: To explore approaches for extracting metrics about authoritativeness in online health communities and how these metrics positively correlate with the quality of the content.

Methods: We designed a metric, called HealthTrust, that estimates the trustworthiness of social media content (eg, blog posts or videos) in a health community. The HealthTrust metric calculates reputation in an online health community based on link analysis. We used the metric to retrieve YouTube videos and channels about diabetes. In two different experiments, health consumers provided 427 ratings of 17 videos and professionals gave 162 ratings of 23 videos. In addition, two professionals reviewed 30 diabetes channels.

Results: HealthTrust may be used for retrieving online videos on diabetes, since it performed better than YouTube Search in most cases. Overall, of 20 potential channels, HealthTrust's filtering allowed only 3 bad channels (15%) versus 8 (40%) on the YouTube list. Misleading and graphic videos (eg, featuring amputations) were more commonly found by YouTube Search than by searches based on HealthTrust. However, some videos from trusted sources had low HealthTrust scores, mostly from general health content providers, and therefore not highly connected in the diabetes community. When comparing video ratings from our reviewers, we found that HealthTrust achieved a positive and statistically significant correlation with professionals (Pearson r_{10} = .65, P = .02) and a trend toward significance with health consumers (r_7 = .65, P = .06) with videos on hemoglobinA_{1c}, but it did not perform as well with diabetic foot videos.

Conclusions: The trust-based metric HealthTrust showed promising results when used to retrieve diabetes content from YouTube. Our research indicates that social network analysis may be used to identify trustworthy social media in health communities.

(J Med Internet Res 2012;14(1):e22) doi:10.2196/jmir.1985

KEYWORDS

Medical informatics; information storage and retrieval; video; online systems; health communication; diabetes

Introduction

The Internet is emerging as one of the main sources of consumer health information [1,2]. Many health authorities, medical associations, hospitals, and patients have published or are publishing online content, including through social media platforms (eg, blogs, YouTube, or Twitter). Kaplan and Haenlein defined social media as consisting of a "set of Web applications, which allows the creation and exchange of user-generated content" [3]. Social media are becoming increasingly mainstream in the health domain [4-6]. For example, there are more than 500 channels on YouTube created by American hospitals, containing thousands of videos [7]. Similarly, the United Kingdom's National Health Service has published more than 500 videos on YouTube [8].

Despite the large volume of good-quality health information available on social media platforms, finding accurate and trustworthy health information can be surprisingly difficult [9-13]. There is a great deal of misinformation, and one often comes across content promoting anorexia or avoiding vaccinations [14,15]. Sometimes bogus health information can become extremely popular and viral (eg, conspiracy theories about vaccination). Therefore, sifting through this to find trustworthy health information remains one of the main challenges faced by health consumers.

In conjunction with the large quantity of information, many health consumers rely on online communities for relevant information. Indeed, online health communities have been found to be very effective in filtering misleading information [16]. Members of online communities have to build their trust gradually, which makes it hard for sources that are not trusted to disseminate misinformation. It is also possible to ask peers about high-quality health information; however, peers are not available all the time and often cannot provide instant feedback.

The objective of this study was to explore approaches for extracting metrics about authoritativeness in online health communities and how these metrics would positively correlate with the quality of the content. An authoritative member of the community (such as the American Diabetes Association) tends to publish or endorse content of better quality than do nonauthoritative members of the community. Using link-based analysis, we extracted a metric (called HealthTrust) about authoritativeness in a health community. We then implemented an algorithm for searching videos and channels based on HealthTrust and tested it with online diabetes content from YouTube.

Background

Outside of the health domain, human experts are rarely used in any scalable fashion for classifying and retrieving webpages. Web information retrieval systems rely on automatic approaches to harvest reputable online resources, mainly based on the analysis of links between pages [17-21]. In Google's PageRank, links from one site to another can be modeled as an endorsement, and they are used to calculate a global rank of all the websites [18]. Another example is the hyperlink-induced topic search (HITS) algorithm [17]. As explained in the following section, HITS is a link analysis algorithm for ranking webpages based on two scores: authoritativeness and hubs. Hubs are essentially webpages that function as directories that have links to authoritative pages. The authorities are webpages that are linked by many of the most representative Webs, so they have a high authoritativeness within the community of Webs. Other algorithms, such as TrustRank, take into account trustworthiness in online communities, aimed at making the search more robust to Web spam [20]. Gou et al explored how to use social network analysis for ranking online videos in a personalized manner [21]. Mislove et al studied the integration of general-purpose social networks with online Web searches [22].

To our knowledge none of those algorithms have been studied in the health domain. One of the main challenges in the health domain is that misleading health information can be very popular (eg, antivaccination videos) and therefore may be paradoxically highly rated and not considered spam by general information retrieval algorithms.

Health consumers need tailored tools to help them find good-quality health social media and websites. A common approach consists in creating quality labels for trustworthy health websites that adhere to a set of guidelines [8,23,24]. Some studies have pointed out cases where those guidelines were not that effective for finding good health information [10,25]. Another difficulty is choosing among dozens of guidelines [23,24]. These guidelines have been combined with automatic approaches that extract certain quality indicators [11,12,26-28] used for online health information retrieval. However, automatic approaches are still not widely used. To our knowledge, none of these projects focus on link-based analysis and trust metrics of health websites, as generic search engines do. In addition, despite the popularity of health videos, we have not come across any project specifically aimed at developing tools to help find relevant health videos.

Methods

In the next subsection, we describe the metric HealthTrust and how it can be integrated to enhance the search of social media content (ie, YouTube diabetes videos). In the subsequent subsections, we describe two studies aimed at evaluating the relationship between the HealthTrust scores of diabetes videos and channels, and their quality as perceived by end users. We designed these experiments to evaluate our hypothesis that HealthTrust's metric can be used to improve the retrieval of health social media. In the first study, we evaluated the use of HealthTrust for filtering diabetes channels from YouTube. In a second study, we evaluated the correlation between HealthTrust scores and ratings of videos about diabetes *A1c testing* and *diabetic foot*.

HealthTrust

According to the Merriam Webster Dictionary, trust is an "assured reliance on the character, ability, strength, or truth of someone or something" [29]. Other related terms, treated as equivalent to *trust*, are *authoritativeness* ("clearly accurate or knowledgeable" [30]) and *reputation* ("overall quality or

character as seen or judged by people in general" [31]). In the Web information retrieval domain, trust has normally been based on the analysis of link structures. A link from one website to another implies an endorsement of the linked website; this approach is very similar to the calculation of impact factors for journals. Trust in the health domain is mainly related to the concept of authoritativeness in terms of the reliability and knowledge of the content creator. There are, however, many additional aspects related to trust such as *appearance* and *impartiality* [32]. Within the scope of this study, we define trust as the "assured reliance on the quality of users and content within an online health community."

As we mentioned in the introduction, online health communities can be effective in filtering out misleading health information [16]. Users disseminating misleading information have a hard job gaining trust within the community. A user creating videos about herbal cures for diabetes will receive less endorsement from the diabetes community than a video created by the American Diabetes Association.

We assume that misleading information will be less endorsed within the health community. Consequently, trustworthiness within the health community will correlate with higher content quality. To compute the trustworthiness of health social media, we designed an algorithm to calculate a metric, called HealthTrust, that estimates the trustworthiness of social media content (eg, blog post or video) in the health community to which it belongs. To evaluate HealthTrust we designed an algorithm for searching online health videos based on that metric.

HealthTrust Metric

Users and content are heavily interconnected in the context of health social media. Figure 1 shows that links between users and content form a graph that models a social network where it is possible to calculate trust-related metrics. Content and users are interconnected and can form a health community with a common interest (eg, diabetes).

HealthTrust (Figure 2) is a metric about trust of content and members of a health community. Trust can be modeled for both users ("I trust this author") and content ("I trust this content"). In fact, your trust in a particular piece of content should be a combination of how much you trust its author and the content itself. Based on these considerations we designed the HealthTrust metric. To calculate HealthTrust a set of steps must be followed: (1) extraction of the community where HealthTrust is going to be applied, (2) calculation of the authoritativeness scores for content and users based on their links, and (3) calculation of HealthTrust scores. Finally, this score can be used for information retrieval purposes as explained in the subsection "HealthTrust for Search."

Figure 1. Example of a health social network.

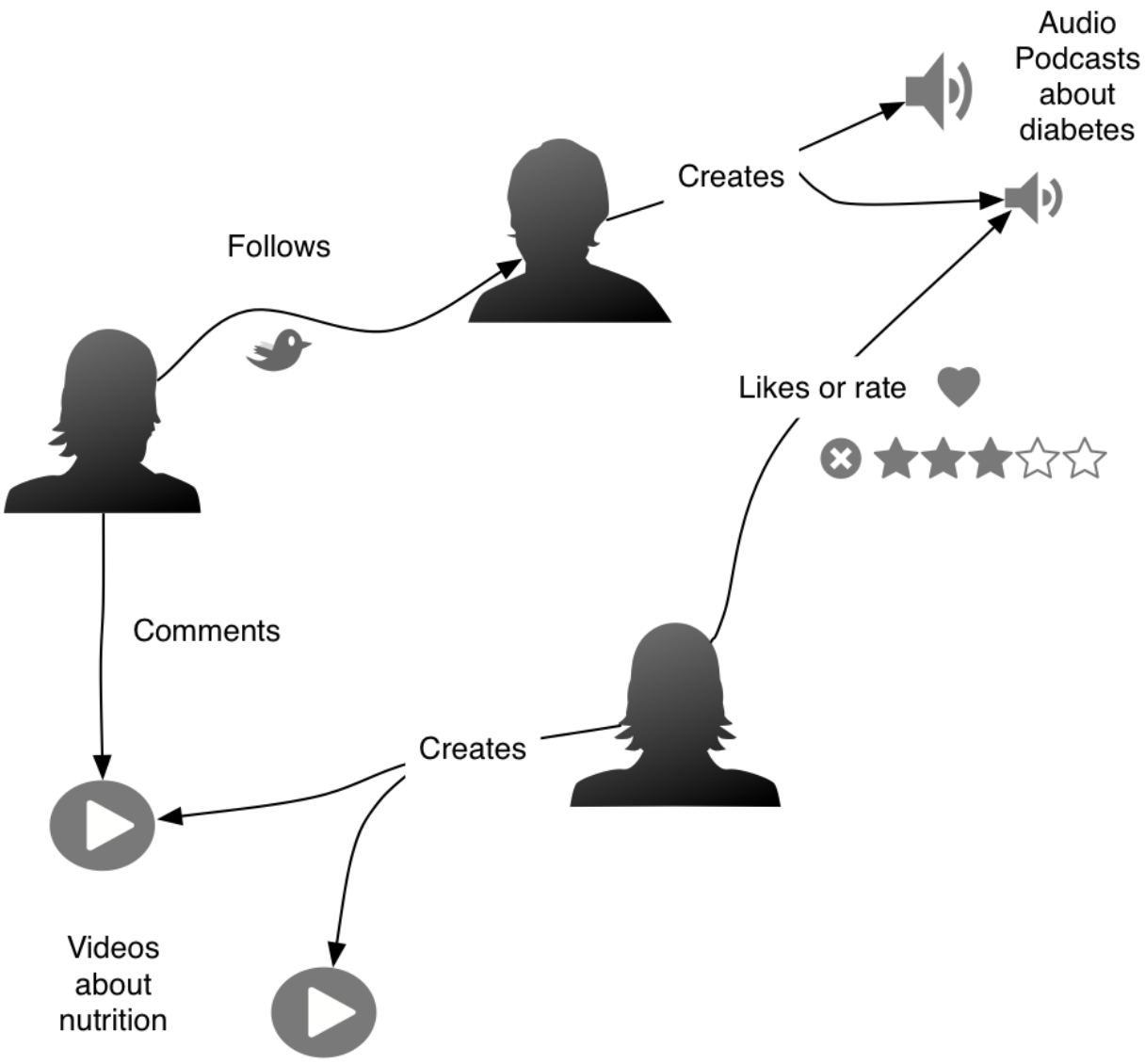


Figure 2. Calculation of the HealthTrust content score.

$$HealthTrust\ (content, community)$$

$$= Authoritativenes\ (content, community)\ x\ (1 - InheritanceFactor)$$

$$+ Authoritativeness\ (author\ of\ content, community)\ x\ InheritanceFactor$$

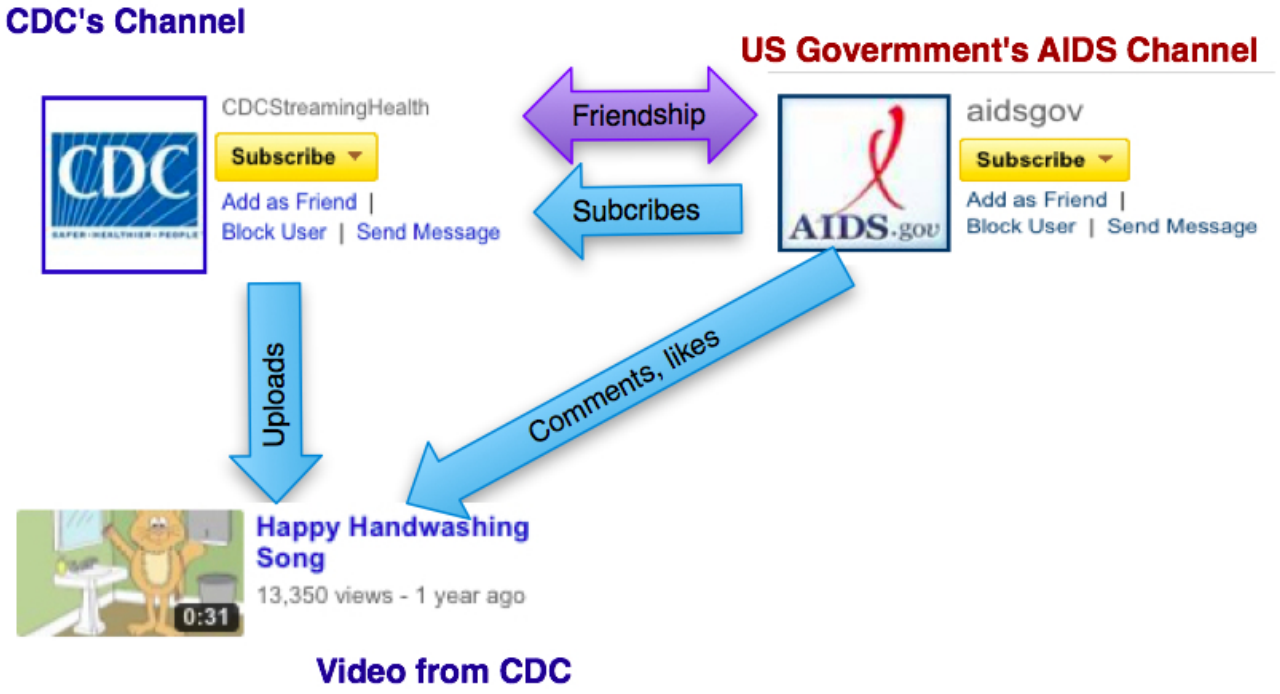
Community Extraction

HealthTrust is applied to only a certain health community. That community can be identified by many different means, such as manual selection of users and heuristic approaches [33]. As explained in the following section, in our study we extracted YouTube users interested in diabetes by using different search queries related to diabetes. Community extraction is a core aspect in HealthTrust, since the metric is not calculating the general authoritativeness of the content but rather the authoritativeness in a particular community. In the case of YouTube in general, MTV videos from rock stars may be more authoritative than videos from health agencies such as the US Centers for Disease Control and Prevention (CDC). On the contrary, with HealthTrust the focus is on intracommunity authoritativeness. For example, in the health community the CDC is far more authoritative than MTV.

For our case study we used the diabetes community on the online video-sharing platform YouTube. As shown on Figure 3, YouTube can be modeled as a social network where users (ie, channels) can build their reputation using different social links (eg, subscriptions, friendships, favorite videos, or comments) [34]. In our study, we took into consideration favorite videos and subscriptions, since these links are the most commonly used by all members of the community.

In our first study we used the YouTube application programming interface (API) to search all the channels that had the keyword *diabetes* and extracted all the accessible information about them (eg, uploads, subscriptions, and favorites). In our second study, we extracted community searches for videos using a set of diabetes-related queries. We extracted all the information about these videos and their authors.

Figure 3. YouTube's social network. CDC = Centers for Disease Control and Prevention.



Authoritativeness Scores

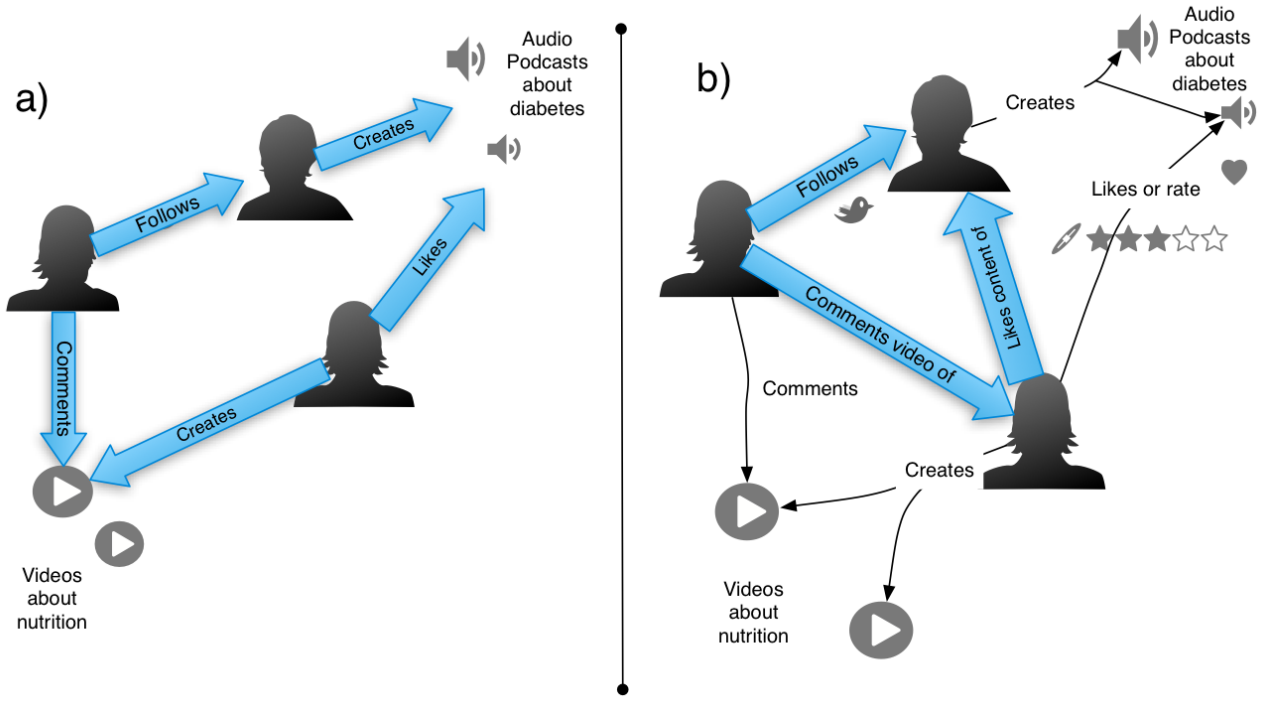
The authoritativeness scores in HealthTrust can be calculated using link-based metrics such as PageRank scores [18] or HITS authoritativeness [17]. As explained in the next section, in our study we used the HITS authoritativeness score. In these algorithms, the links between websites are used to model a bidirectional graph with incoming and outgoing links. A recursive algorithm is used to score the reputation of a website based on the incoming links, since an incoming link is considered an endorsement of the linked website. The HITS algorithm considers two types of nodes: authorities and hubs. The hubs are the nodes that tend to link to the most authoritative webpages. The authoritative scores in HITS are calculated based on the incoming links from hubs.

The authoritativeness of content and users are calculated as follows. First, the authoritativeness of content (Figure 4, left) is calculated based on the links between all users and content. Both content and users are considered nodes. Second, the authoritativeness of users (Figure 4, right) is calculated based on the links between all users, which are the only nodes. If a user likes or favors content from another user, this is considered as a link between the users.

In our study we used the Java Universal Network/Graph (JUNG) API [35] to calculate the HITS authoritativeness values of users (ie, channels) and content (videos) as follows. First, for the authoritativeness of users, we created a graph where the nodes were the channels and the edges were their subscriptions (channel X subscribed to channel Y) and favorites (channel X subscribed to video of channel Y). Then, that graph was used to calculate the HITS authoritativeness values of the channels. Second, for the authoritativeness scores of videos, we considered videos and channels to be nodes and the edges to be favorites (channel X subscribed to video Z) and subscriptions (channel X subscribed to video of channel Y). That graph was used to calculate the HITS authoritativeness values of the videos.

The authoritativeness values for content and users are calculated independently as shown in the Figure 4. Therefore, to combine them it is necessary to normalize their ranges—for example, in our study we normalized the authoritativeness scores of videos and users between 0 and 1.

Figure 4. Links (in blue) used to calculate authoritativeness of users (left) and content (right). Diagram based on Figure 1.



Calculation of HealthTrust

The HealthTrust score of a particular piece of content (eg, video or blog post) is the weighted combination of the normalized authoritativeness scores of content. The weighted combination is based on the InheritanceFactor. The weighted approach is designed to allow part of the trustworthiness to be inherited by the content from its author. Thus, new content from a trusted author will have a higher HealthTrust score than new content from an untrustworthy author. To give a high weight to the InheritanceFactor implies that the authoritativeness of the author is very important. For example, a video from the CDC will have implicit authoritativeness even if it is new and has never been rated or linked. An InheritanceFactor of 0 implies that there is no inheritance transfer of trust from the authors to their content, so all the authoritativeness is based on the video's score.

As Figure 5 shows, in the video study authoritativeness scores were combined with an InheritanceFactor of 0.7, meaning that the HITS authoritative value of videos weighed 30% and the author's authoritativeness 70%. We decided on these values after testing with several queries (not used in our evaluation) in a previous data set. We observed that there were many new high-quality videos without links to them, so a lower value for the InheritanceFactor would have decreased their HealthTrust despite being from a trusted content provider.

Figure 5. HealthTrust calculation for diabetes videos from YouTube.

$$HealthTrust\ (video\ v, diabetes\ comunity)$$
$$= Authoritativenes\ (video, community)\ x\ (1 - 0.7)$$
$$+ Authoritativeness\ (author\ of\ video, community)\ x\ 0.7$$

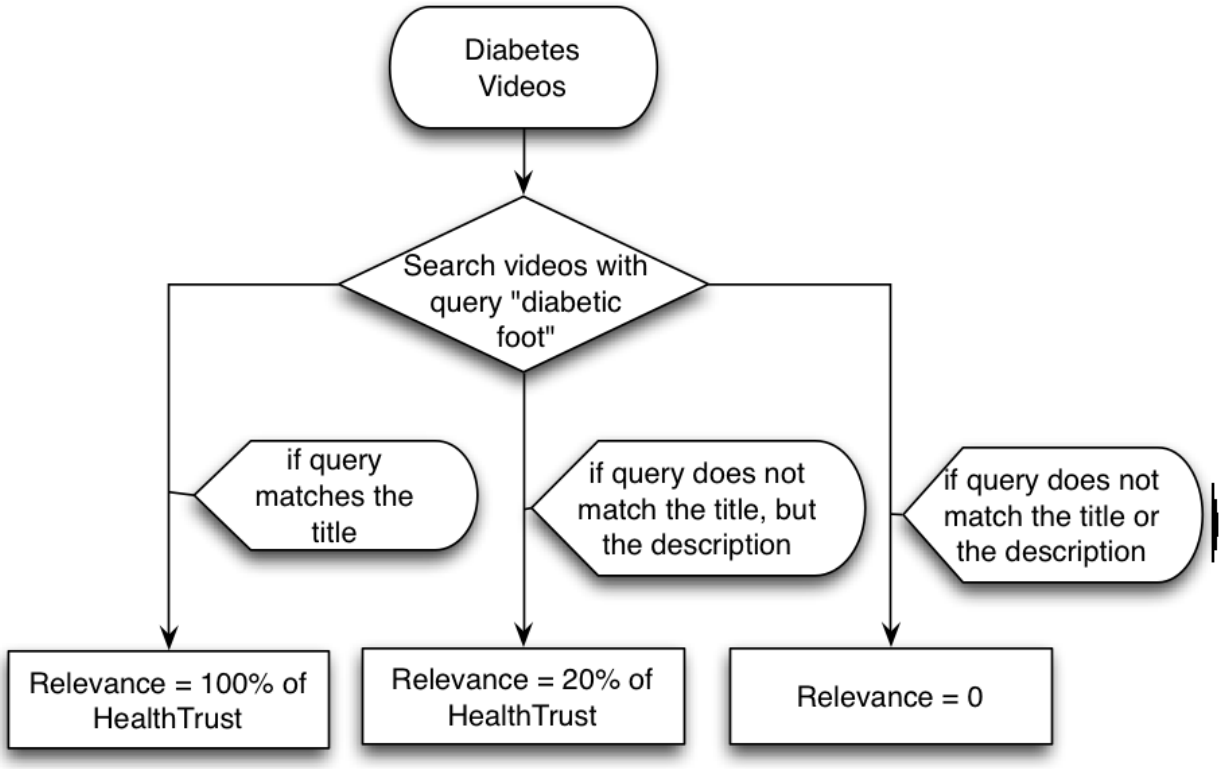
HealthTrust for Search

We believe that HealthTrust can be used to enhance the retrieval of content within health communities. To evaluate that possibility we designed a search algorithm that combines query matching with HealthTrust. Our search algorithm is based on combining two scores: (1) relevance of the content to the search query, and (2) HealthTrust. Relevance can be calculated using simple query matching (eg, the content contains the query in its title or in the description).

We implemented a search algorithm based on HealthTrust to study whether that metric may be use to retrieve diabetes videos. As shown in Figure 6, our search algorithm combined the syntactic query match with the HealthTrust values. If the query matched the video's title the relevance was computed as 100% of the video's HealthTrust score. If the query only matched the description, the relevance was computed as 20% of the video's HealthTrust score. We decided on these values after observing the quality of video metadata. In particular, we observed that titles are very important to infer the relevance of videos, since descriptions and tags tend to be very heterogeneous (eg, due to tag spamming). In a previous study, we also found that the

quality of comments on YouTube health videos can be very heterogeneous [36].

Figure 6. Relevance calculation for HealthTrust-based search.



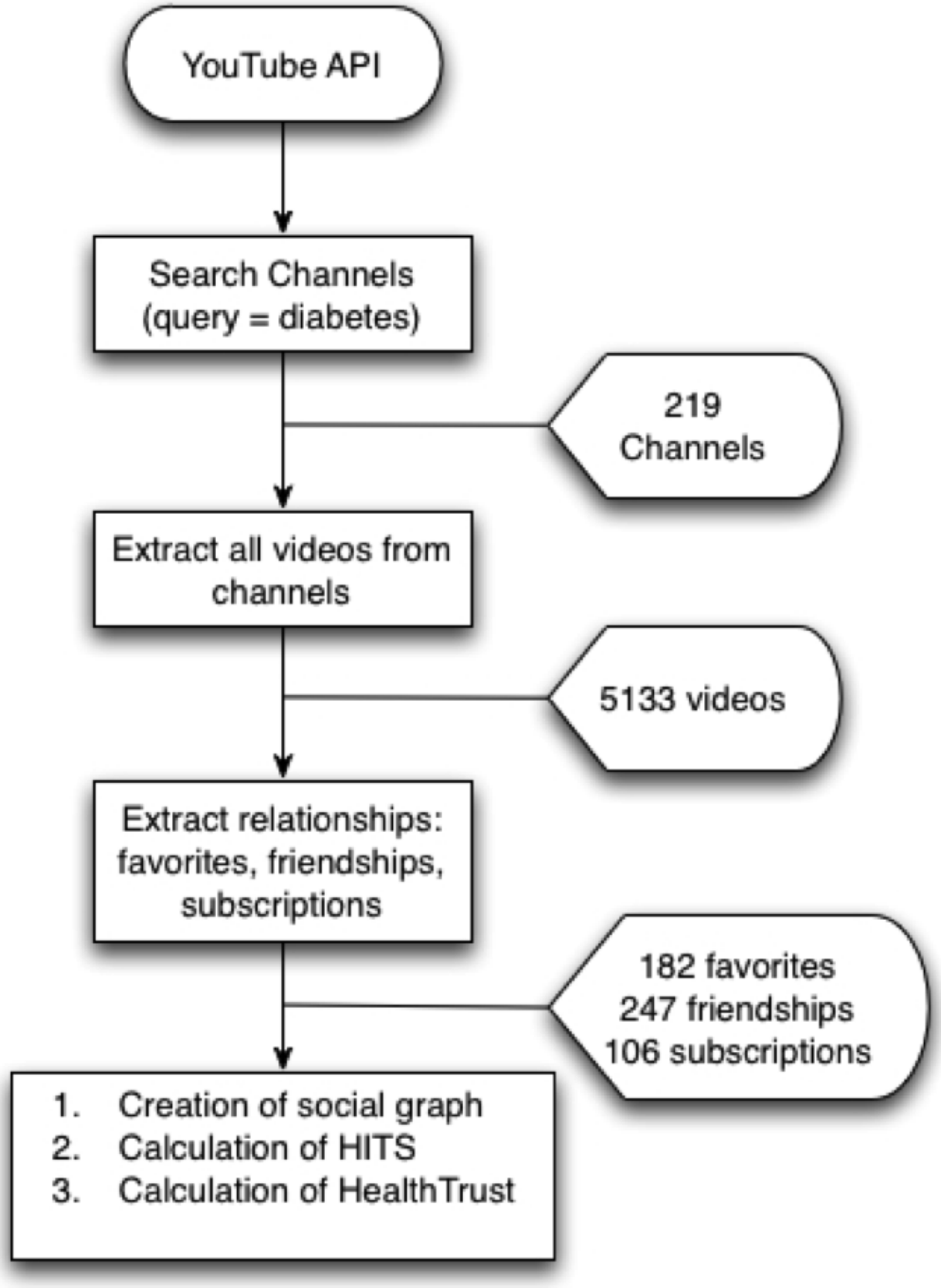
Study: Diabetes Channels and HealthTrust

As described in a previous report [37], in May 2010, we performed a study to evaluate the feasibility of using social network analysis to filter YouTube diabetes channels. The objective of this study was to test whether the authoritativeness values of the diabetes channels in YouTube are related to their quality.

Data Collection

Figure 7 describes how we extracted 5133 videos, 219 channels, 182 favorites, and 247 friendships about diabetes from YouTube using the YouTube API. We searched channels with the query *diabetes* and extracted their information (links, videos, descriptions, etc) to calculate their HealthTrust scores, which corresponded to the authoritativeness values of the channels, since we did not take videos into consideration in this study.

Figure 7. Data extraction in the study of diabetes channels and HealthTrust. API = application programming interface; HITS = hyperlink-induced topic search.



Recruitment and Ratings

Two health care professionals rated channels from a list containing the top 20 diabetes channels retrieved by YouTube and HealthTrust's top 20 channels. The reviewers received a list with all the channels alphabetically ordered and were asked to respond with "yes" or "no" to whether they would recommend the diabetes channel to a patient with diabetes.

The interrated agreement score based on Cohen kappa [38] was calculated using the statistical framework R [39] and resulted in good agreement (.61).

Data Analysis

We evaluated the results using the precision at K metric [40], with K being the top-ranked retrieved results. This technique is used widely to evaluate Web search engines, since users tend to use only the top search results. We also evaluated the results with the metric discounted cumulative gain (DCG). DCG is commonly used to evaluate ranked lists of Web search results taking into account the position of the retrieved results [41]. The relevance *gain* decreases logarithmically based on the position of the retrieved results.

Search Study: Diabetes Videos and HealthTrust

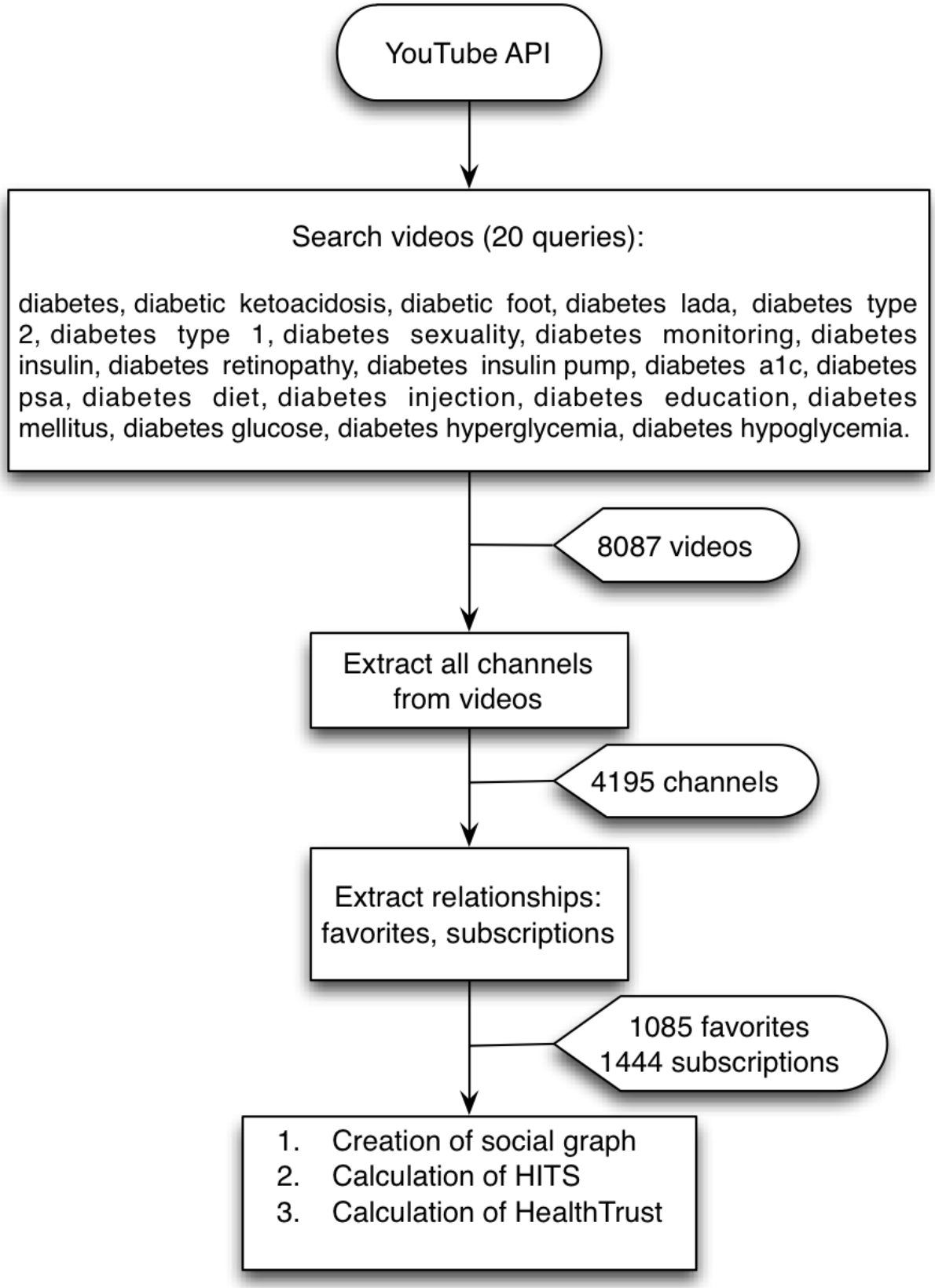
Data Collection

In April 2011, we collected from YouTube 8087 diabetes videos using the search API with 20 different queries (*diabetic foot*, *diabetes*, *diabetes ketoacidosis*, etc) as explained in Figure 8. We also extracted all the available information about channels, subscriptions, and favorites. Finally, we calculated the HealthTrust scores for videos and channels.

Although our dataset contained videos found by different queries, we evaluated videos from only two queries in order to increase the number of responses per video. We limited our study to the evaluation of searches about two information needs that are important for most people affected by diabetes: diabetes foot issues and hemoglobin A_{1c} (glycated hemoglobin) testing. Diabetes foot problems are very common among people with diabetes and require a lot of attention to avoid very serious complications that can lead to amputation. Diabetes hemoglobin A_{1c} testing is a very common laboratory test to evaluate how well the diabetes is managed.

Most of the responders rated different videos, since there were four different lists and some of the surveys were not completely filled out. Therefore, there was not enough data to calculate a meaningful interannotator agreement score in this study. For each type of responder (professionals and consumers), we aggregated the ratings of the different videos and calculated the average rating values.

Figure 8. Data extraction on the search study on diabetes videos and HealthTrust. API = application programming interface; HITS = hyperlink-induced topic search.



Recruitment

After extracting the dataset of diabetes videos we recruited professionals and health consumers to evaluate the results. The recruitment took place between April 25 and June 14, 2011.

We recruited health care professional reviewers using a snowball approach, where invitations were sent to professional mailing lists. We collected 82 informed consents, and 27 video surveys were completed (2 surveys were removed due to the lack of information about the profession of the respondents). In total, professionals provided 162 ratings of 23 videos.

We recruited health consumers from the online diabetes community TuDiabetes.org, which has more than 20,000 members. Information about the study was posted on the community's main blog and in their mailing list (about 10,000 subscribers). We received 178 informed consents, and 73 surveys were partially or completely filled in. In total, consumers provided 427 ratings of 17 videos. A donation of US $5 per survey was given to the Diabetes Hands Foundation, which runs the online community.

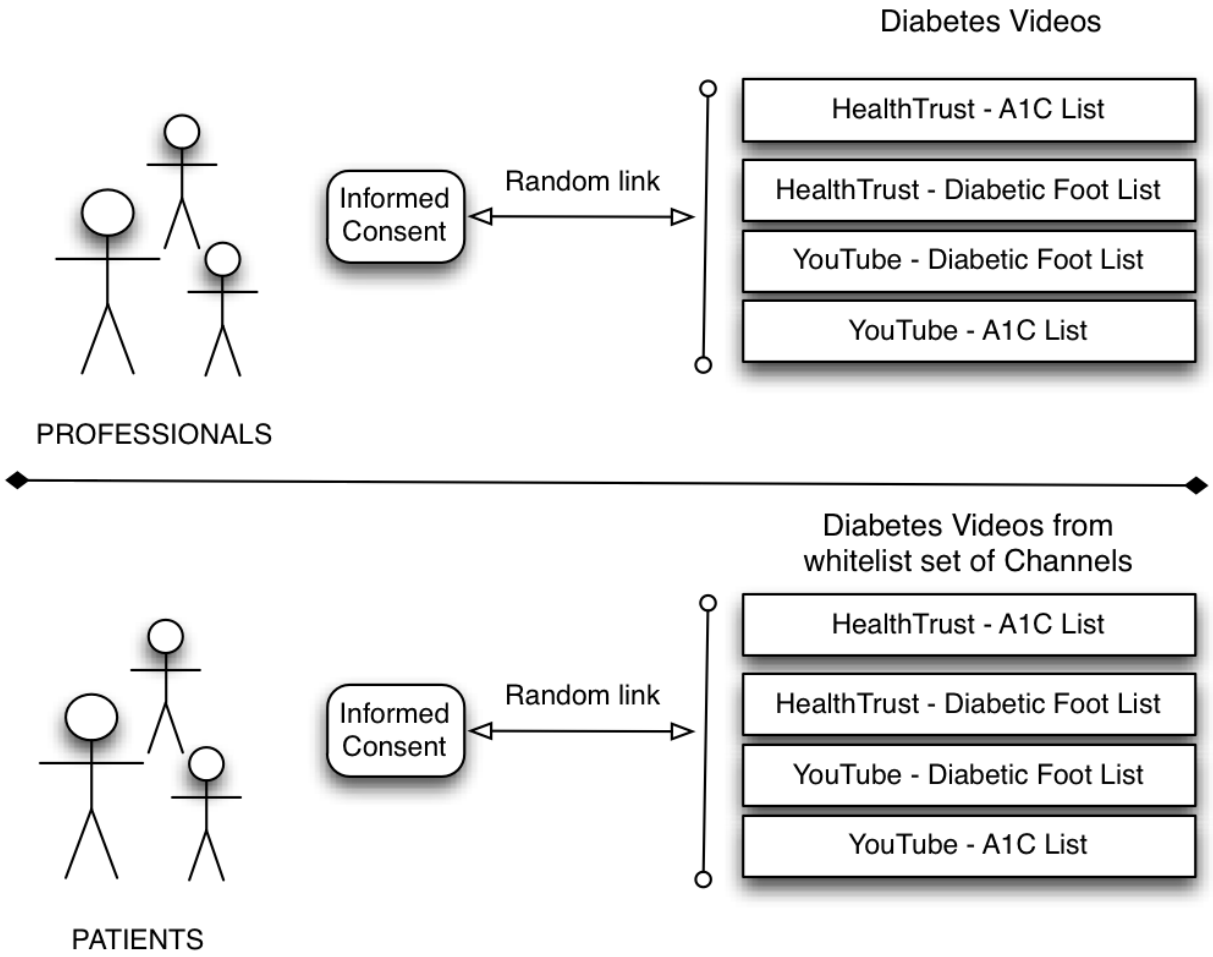
Video Surveys

We evaluated the top 7 video search results for the queries *diabetic foot* and *diabetes A1c* using both HealthTrust and YouTube search (ordered by relevance). As depicted in Figure 9, after providing informed consent in a Web form the respondents were randomly assigned to a survey with videos to review. Respondents were not informed about the algorithm used to select the videos.

Professional reviewers were assigned to one of four different surveys: two about diabetic foot (one based on YouTube and the other on HealthTrust) and two for hemoglobin A_{1c} testing. The two lists about diabetic foot for professionals contained 11 videos and the lists about hemoglobin A_{1c} testing contained 12 videos. Health consumers were assigned to lists for the same queries, but the listed videos were limited to those published by a whitelist of trusted sources. The main reason for using a whitelist was to avoid showing misleading and disturbing videos to consumers (eg, videos featuring amputations). These lists for health consumers contained a total of 17 videos, 8 about diabetic foot and 9 about hemoglobin A_{1c} testing.

Professionals and health consumers were asked to respond to the following questions about the videos with a Likert scale (eg, from strongly agree to strongly disagree): "Would you recommend this video to a patient with diabetes and questions about diabetic foot?" (question for professionals for a video from the list about diabetic foot); and "Do you like this video about diabetic foot?" (question for health consumers for a video from the list about diabetic foot).

Figure 9. Process of obtaining informed consent from health care professionals and health consumers, and survey allocation.



Data Analysis

We evaluated the retrieved results using the metrics precision at K [40] and DCG [41]. However, we did not calculate either of these metrics for the health consumers, as they had a prefiltered dataset.

In addition, we used the Pearson correlation [42] to study the correlation between the HealthTrust scores and the average ratings. Pearson correlation is commonly used to study linear dependence between two variables, and the correlation coefficient ranges from −1 to 1. The Pearson correlation was calculated using the psych package of statistical framework R [39].

Results

Study of Diabetes Channels

The first study was designed to evaluate the feasibility of using the HealthTrust metric to filter YouTube diabetes channels (aka users). We studied precision at K (K = 5, K = 10, and K = 20) in the top 20 diabetes channels retrieved by the YouTube- and HealthTrust-based searches.

We proposed two possible scenarios for considering a channel relevant: (1) both reviewers recommended the channel and, (2) at least one reviewer recommended the channel. Table 1 shows that the search based on HealthTrust scores performed better than YouTube search in all cases and was only equally good for precision at K = 5 and for channels recommended by both reviewers. The DCG evaluation (Table 2) also resulted in better scores for HealthTrust than for YouTube searches.

Table 1. Evaluation of the top 20 diabetes channels by precision at K metric

Recommended by/precision at K	Both reviewers				At least one reviewer			
	YouTube		HealthTrust		YouTube		HealthTrust	
	n	%	n	%	n	%	n	%
K = 5	4	80%	4	80%	4	80%	5	100%
K = 10	6	60%	7	70%	7	70%	9	90%
K = 20	10	50%	13	65%	12	60%	17	85%

Table 2. Evaluation of the top 20 diabetes channels by discounted cumulative gain (DCG) metric

Recommended by/DCGi[a]	Both reviewers		At least one reviewer	
	YouTube	HealthTrust	YouTube	HealthTrust
i = 5	2.9	3.1	2.9	3.6
i = 10	3.6	4.1	4	4.9
i = 20	4.6	5.7	5.3	7

[a] i = number of retrieved videos.

To consider and analyze the capacity of the algorithms to filter out bad content or spam, we considered a channel to be misleading if none of the reviewers recommended it. HealthTrust's approach performed quite well, filtering out bad channels. For K = 20, HealthTrust's list had only 3 bad channels (15%) versus 8 (40%) on the YouTube list. In the top 10 channels, HealthTrust had only 1 bad channel (10%) versus 3 (30%) for YouTube. Within the top 5 channels, all HealthTrust's channels were recommended by at least one reviewer, one more than YouTube.

In the YouTube top 20, some channels featured commercials of diabetes products (eg, testing supplies), several were about a famous diabetic singer (Jonas), and one channel was in Dutch (even though we restricted the search to English in the API). The YouTube list also contained some channels with the word diabetes in its name, but most of the videos were not related to diabetes.

The HealthTrust list did not contain any channels with advertising, but it did have some channels from e-patients with very heterogeneous quality. Surprisingly, some diabetes channels run by public authorities, such as the Juvenile Diabetes Research Foundation, were not highly ranked in HealthTrust. The most logical explanation for this is that some relevant channels do not encourage social interactions (eg, friendships or subscriptions), and this less-connected nature may decrease their rankings.

Study of Diabetes Videos

In the second study, we explored how the HealthTrust metric can be used to retrieve diabetes videos and also the correlation between HealthTrust and the video's ratings.

HealthTrust Search Evaluation

We calculated precision at K for the list created for professionals to evaluate the performance of the search algorithm. However, we did not study precision at K for consumers, since the dataset was prefiltered.

Precision at K (K = 3, K = 7) for the professionals' lists was considered as a video rating equal to or greater than 3.5 (values range from 1 to 5). As shown in Table 3, precision was better

in HealthTrust for both the diabetes A_{1c} and diabetic foot lists. In the case of diabetic foot, the YouTube list precision was below 50% for both the top 7 and the top 3. The HealthTrust-based search also performed better based on the DCG metric (Table 4).

Table 3. Precision at K for videos evaluated by professionals retrieved by HealthTrust and YouTube

Precision at K	Hemoglobin A_{1c}				Diabetic foot			
	YouTube		HealthTrust		YouTube		HealthTrust	
	n	%	n	%	n	%	n	%
K = 3	2	66%	3	100%	1	33%	2	66%
K = 7	4	57%	5	70%	3	43%	4	57%

Table 4. Discounted cumulative gain (DCG) for videos evaluated by professionals retrieved by HealthTrust and YouTube

DCGi[a]	Hemoglobin A_{1c}		Diabetic foot	
	YouTube	HealthTrust	YouTube	HealthTrust
i = 3	1.6	2.6	1	2
i = 7	2,6	3.4	1.9	2.9

[a] i = number of retrieved videos.

HealthTrust and Rating Correlation

The study of the correlation between HealthTrust score and average rating was used to determine whether our trustworthiness score had a positive impact on the ratings.

For both professionals and consumers, we created two subsets with the videos of each topic (hemoglobin A_{1c} testing and diabetic foot). We normalized the average ratings of the videos between 0 and 1 for the subset with the videos about hemoglobin A_{1c} testing and the subset about diabetic foot. Similarly, we normalized the HealthTrust scores within each subset. As shown in Table 5, we compared average ratings and HealthTrust scores for each subset using the Pearson correlation (alpha = .05).

Table 5. Pearson correlation between ratings and HealthTrust

	Hemoglobin A_{1c}		Diabetic foot	
	Pearson *r*	*P* value	Pearson *r*	*P* value
Professionals	r_{10} = .646	.02	r_9 = .275	.41
Health consumers	r_7 = .649	.06	r_6 = –.019	.96

In the case of the hemoglobin A_{1c} videos, we found a positive and statistically significant correlation for the professionals' subset (Pearson r_{10} = .646, P = .02). This correlation was weaker with the health consumers but still close to significance levels (r_7 = .649, P = .06). In the case of the diabetic foot videos, we did not find a statistically significant result in any of the subsets (professionals and consumers).

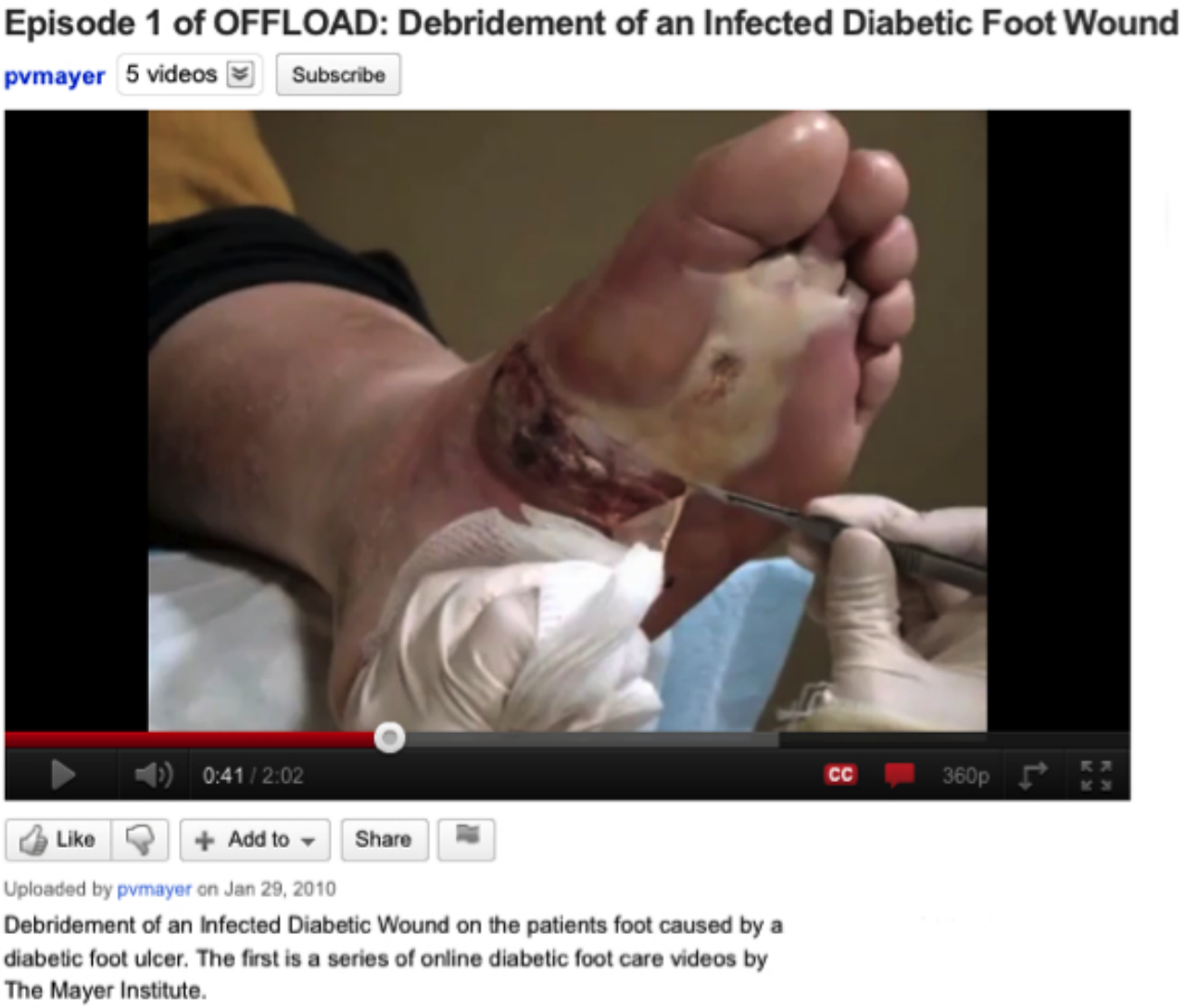
Discussion

HealthTrust Metric Performance

Our results suggest that social network analysis may be used to gather information about the quality of health information. The retrieval of diabetes videos and channels based on the HealthTrust metric performed reasonably well, compared with the YouTube search. In nearly all cases, the precision of the lists retrieved using HealthTrust was better than those retrieved using YouTube. Precision is very important, since in the health-irrelevant content can be potentially very negative (see Figure 10). It is quite significant that the performance of HealthTrust was equal to or better than that of the YouTube search, considering that YouTube has access to all the metadata about videos and users, while HealthTrust has limited access via its API; for example, some channels restrict access to information about their links (eg, subscriptions).

It is difficult to identify the exact differences between the YouTube and HealthTrust searches, since YouTube has not published its search algorithm, despite having published its recommended algorithm [43]. However, we expect YouTube's search engine to be based on link analysis, as are most search engines. The main difference between traditional Web search engines and our approach is that we strengthened the tightly knit community effect [19], as with diabetes; traditionally, Web search engines try to reduce the influence of those communities to raise general public satisfaction. Consequently, funny or controversial videos are more popular among the general YouTube community and are therefore more highly rated. These videos lose prominence using HealthTrust. In fact, the search based on the HealthTrust metric performed better than YouTube in filtering out misleading videos (eg, herbal cures or advertisements). The HealthTrust algorithm estimates health-related trust and not general trust on YouTube.

Figure 10. Highly ranked YouTube video about diabetic foot featuring an infected wound.



HealthTrust Weaknesses

Some good videos from trusted sources, such as public health authorities, gained relatively low HealthTrust values. The best explanation for the algorithm's behavior is that creators of those videos had fewer connections in the diabetes community. Some of these sources belonged to a more generic health community (eg, CDCStreamingHealth's Channel) and therefore had weaker ties with the diabetes community. Also, some trusted sources do not create links with other users (eg, through friendships, subscriptions, or comments). This lack of connectivity leads to lower scores in HealthTrust. As part of our future work, we will design an enhanced version of HealthTrust that calculates trustworthiness values within several health communities.

Many factors influence the perceived quality of a video beyond trustworthiness and authoritativeness. Therefore, it is not surprising that we did not find statistically significant correlations in all cases. Personal taste and preferences play a major role. For example, the video *O is for outrage – Type 1 diabetes* (Figure 11) was given a higher rating by health consumers (average of 4.2) than by professionals (average of 2.75). *O is for outrage* is a video appealing to emotional aspects to raise awareness; it is very engaging to the online diabetes community. However, this particular video is less informative, which may explain why professionals rated it lower. Consequently, a generic quality indicator such as HealthTrust cannot always satisfy everybody.

There were videos from certain channels with quite different average ratings but the same HealthTrust scores. In such cases, the videos had no links (favorites) but inherited the HealthTrust scores from their channels. An example of this problem is shown in the following two videos from the diabetic foot list for consumers: (1) *Baseball great Ron Santo & Diabetes--INCREDIBLE Story*, and (2) *Miami Ink's Darren Brass: Tattoos and Diabetes*. Both videos have the same HealthTrust score, as both are from the same diabetes channel, dLifedotcom. However, the *Miami Ink* video was less appealing to health consumers. In this case, link analysis was not enough to distinguish the quality between the two videos. The only way to solve this problem is to analyze more data about the video (eg, semantic analysis or ratings).

Figure 11. Screenshot from the video *O is for outrage – Type 1 diabetes*.

Limitations

In both experiments, some videos or channels were deleted while we were conducting the experiment. In the case of the channel study, two were removed, and in the second study some videos gained lower ratings because they were made private by their authors. It is unlikely that this problem biased our study, since it affected a small sample and it affected all the algorithms equally.

To be able to generalize our findings, we will have to perform large-scale studies with more queries, reviewers, and videos. Our survey-based evaluation approach is merely an approximation of the real context of health consumers' search for information. Survey-based evaluation of online videos is very time consuming, as most videos last several minutes. It was necessary to watch around 30 minutes of videos to complete our surveys. Hence, to generalize our findings we are implementing a video portal to capture more data for evaluation within the real user context. The video portal will also need to address the continuous changes in the structure of online communities (eg, reputation changes over time). A possible solution for the dynamic nature of online communities may be periodic calculation of HealthTrust.

Moreover, it remains to be seen whether our approach will work in health domains where there is a large community of users promoting misleading information. For example, there are communities promoting anorexia as a lifestyle [15] or against vaccination [14]. Pro-anorexia users will tend to link and endorse misleading information; thus, if HealthTrust is to be used to retrieve trustworthy content about anorexia it must be able to avoid pro-anorexia subcommunities.

Our current study is limited to online health videos; therefore, we will need to replicate our study with other types of social media in order to generalize our findings. We believe that the metric HealthTrust can be applied to any type of linked health community where users are interconnected via follows, friendships, and favorite content. However, experiments will need to be performed to evaluate the algorithm, since each type of community may have a different structure and dynamics.

Our study is limited to automatic approaches for extracting trust-based metrics and the feasibility of using these metrics to retrieve health videos. More research will be needed to test how

to combine HealthTrust with manual selection of social media by human experts. HealthTrust can be very useful to automatically identify the most trusted sources within the diabetes community. However, some trustworthy providers have very good content but have not gained enough trust within the online community.

Conclusions

Every day, millions of health consumers search for health information on social platforms such as YouTube, and retrieving accurate information from trusted sources can often be difficult. There is an unsatisfied need for new information retrieval tools to help health consumers find trustworthy and relevant health information within social media.

In this paper we present a new metric, called HealthTrust, to infer information about the trustworthiness of social media content within a health community. We tested the feasibility of using HealthTrust for retrieving videos from the diabetes community on YouTube. Based on our evaluation with health consumers and professionals, the search of diabetes content based on the HealthTrust metric performed better than YouTube in nearly all the tested cases. However, a larger study is needed to validate our results in a health portal in order to test the metric in a live setting.

Despite the limitations of our study, we conclude that, to apply social network analysis to retrieving health information, social media may be used to develop tools that will ultimately help find relevant and trustworthy information. Social network analysis could also be used to reinforce other approaches to health information retrieval such as quality labels and manual review of content.

Acknowledgments

We would like to thank our colleagues for their very useful comments, which helped to improve this paper, especially Prof. Joseph A Kostan from the GroupLens group in the Department of Computer Science and Engineering at the University of Minnesota. In addition, we would like to thank the Diabetes Hands Foundation for their support in the recruitment of people with diabetes.

This project belongs to the Tromsø Telemedicine Laboratory cofunded by the Research Council of Norway, project 174934.

Authors' Contributions

LFL was involved in all the research and manuscript preparation.
RK reviewed the manuscript and advised LFL with regard to the algorithm aspects of the study.
GBM participated in the study design, institutional review board application, and recruitment, and reviewed the manuscript.

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 2012-01-26] [WebCite Cache ID 64zCRIRq5]
2. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
3. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. Business Horizons 2010;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
4. Lau AY, Siek KA, Fernandez-Luque L, Tange H, Chhanabhai P, Li SY, et al. The role of social media for patients and consumer health: contribution of the IMIA Consumer Health Informatics Working Group. Yearb Med Inform 2011;6(1):131-138. [Medline: 21938338]
5. Fox S, Jones S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2009 Jun 11. URL: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2010-12-07] [WebCite Cache ID 5uno4xyeW]
6. Paton C, Bamidis PD, Eysenbach G, Hansen M, Cabrer M. Experience in the use of social media in medical and health education: contribution of the IMIA Social Media Working Group. Yearb Med Inform 2011;6(1):21-29. [Medline: 21938320]
7. Bennett E. Found in Cache. 2011 Oct 9. Hospital Social Network List URL: http://ebennett.org/hsnl/ [accessed 2011-11-06] [WebCite Cache ID 62zskSpvI]
8. YouTube. 2011. NHSChoices Channel. Sep 12 URL: http://www.youtube.com/user/NHSChoices [accessed 2011-11-06] [WebCite Cache ID 62zsQQ4cu]
9. Greenberg L, D'Andrea G, Lorence D. Setting the public agenda for online health search: a white paper and action agenda. J Med Internet Res 2004 Jun 8;6(2):e18 [FREE Full text] [doi: 10.2196/jmir.6.2.e18] [Medline: 15249267]
10. 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 May;287(20):2691-2700 [FREE Full text] [Medline: 12020305]
11. Bernstam EV, Walji MF, Sagaram S, Sagaram D, Johnson CW, Meric-Bernstam F. Commonly cited website quality criteria are not effective at identifying inaccurate online information about breast cancer. Cancer 2008 Mar 15;112(6):1206-1213 [FREE Full text] [doi: 10.1002/cncr.23308] [Medline: 18266210]

12. Mayer MA, Darmoni SJ, Fiene M, Köhler C, Roth-Berghofer TR, Eysenbach G. MedCIRCLE: collaboration for Internet rating, certification, labelling and evaluation of health information on the World-Wide-Web. Stud Health Technol Inform 2003;95:667-672. [Medline: 14664064]
13. Mayer MA, Karkaletsis V, Stamatakis K, Leis A, Villarroel D, Thomeczek C, et al. MedIEQ-Quality labelling of medical web content using multilingual information extraction. Stud Health Technol Inform 2006;121:183-190. [Medline: 17095816]
14. Ache KA, Wallace LS. Human papillomavirus vaccination coverage on YouTube. Am J Prev Med 2008 Oct;35(4):389-392. [doi: 10.1016/j.amepre.2008.06.029] [Medline: 18675530]
15. Norris ML, Boydell KM, Pinhas L, Katzman DK. Ana and the Internet: a review of pro-anorexia websites. Int J Eat Disord 2006 Sep;39(6):443-447. [doi: 10.1002/eat.20305] [Medline: 16721839]
16. Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. BMJ 2006 Apr 22;332(7547):939-942 [FREE Full text] [doi: 10.1136/bmj.38753.524201.7C] [Medline: 16513686]
17. Kleinberg JM. Authoritative sources in a hyperlinked environment. J ACM 1999;46(5):604-632. [doi: 10.1145/324133.324140]
18. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. Comput Netw ISDN Syst 1998;30(1-7):107-117. [doi: 10.1016/S0169-7552(98)00110-X]
19. Lempel R, Moran S. The stochastic approach for link-structure analysis (SALSA) and the TKC effect. Comput Netw 2000;33(1-6):387-401. [doi: 10.1016/S1389-1286(00)00034-7]
20. Gyongyi Z, Garcia-Molina H, Pedersen J. Combating web spam with trustrank. In: Proceedings. 2004 Presented at: Thirtieth international Conference on Very Large Data Bases; Aug 31-Sept 3, 2004; Toronto, ON, Canada URL: http://portal.acm.org/citation.cfm?id=1316689.1316740
21. Gou L, Chen HH, Kim JH, Zhang L, Lee Giles C. SNDocRank: document ranking based on social networks. In: Proceedings. 2010 Presented at: 19th international Conference on World wide web (WWW '10); Apri 27-30, 2010; Raleigh, NC, USA. [doi: 10.1145/1772690.1772825]
22. Mislove A, Marcon M, Gummadi KP, Druschel P, Bhattacharjee B. Measurement and analysis of online social networks. In: Proceedings. 2007 Presented at: 7th ACM SIGCOMM Conference on Internet Measurement (IMC '07); Oct 24-26, 2007; San Diego, CA, USA. [doi: 10.1145/1298306.1298311]
23. 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 [FREE Full text] [Medline: 11884320]
24. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? JAMA 1998 Feb 25;279(8):611-614 [FREE Full text] [Medline: 9486757]
25. Walji M, Sagaram S, Sagaram D, Meric-Bernstam F, Johnson C, Mirza NQ, et al. Efficacy of quality criteria to identify potentially harmful information: a cross-sectional survey of complementary and alternative medicine web sites. J Med Internet Res 2004 Jun 29;6(2):e21 [FREE Full text] [doi: 10.2196/jmir.6.2.e21] [Medline: 15249270]
26. M-Eco Consortium. meco-project.eu. 2010. M-Eco: Medical EcoSystem URL: http://meco-project.eu/ [accessed 2011-11-06] [WebCite Cache ID 62ztB4MY4]
27. Khresmoi. 2010. Khresmoi: Medical Information Analysis & Retrieval URL: http://www.khresmoi.eu/ [accessed 2011-11-06] [WebCite Cache ID 62ztdsCrb]
28. Mayer MA, Karampiperis P, Kukurikos A, Karkaletsis V, Stamatakis K, Villarroel D, et al. Applying Semantic Web technologies to improve the retrieval, credibility and use of health-related web resources. Health Informatics J 2011 Jun;17(2):95-115. [doi: 10.1177/1460458211405004] [Medline: 21712354]
29. Merriam-Webster Inc. 2012. Definition of Trust URL: http://www.merriam-webster.com/dictionary/trust [accessed 2011-11-06] [WebCite Cache ID 62zttQS1q]
30. Merriam-Webster Inc. 2012. Definition of Authoritative URL: http://www.merriam-webster.com/dictionary/authoritative [accessed 2011-11-06] [WebCite Cache ID 62ztwFea0]
31. Merriam-Webster Inc. 2012. Definition of Reputation URL: http://www.merriam-webster.com/dictionary/reputation [accessed 2011-11-06] [WebCite Cache ID 62ztyjbVq]
32. Harris PR, Sillence E, Briggs P. Perceived threat and corroboration: key factors that improve a predictive model of trust in internet-based health information and advice. J Med Internet Res 2011;13(3):e51 [FREE Full text] [doi: 10.2196/jmir.1821] [Medline: 21795237]
33. Chin A, Chignell M. Automatic detection of cohesive subgroups within social hypertext: a heuristic approach. N Rev Hypermed Multimed 2008;14(1):121-143. [doi: 10.1080/13614560802357180]
34. Han YS, Kim L, Cha JW. Evaluation of user reputation on YouTube. Lect Notes Comput Sci 2009;5621:346-353. [doi: 10.1007/978-3-642-02774-1_38]
35. JUNG Framework Development Team. Sourceforge.net. JUNG: Java Universal Network/Graph Framework URL: http://jung.sourceforge.net/ [accessed 2011-11-06] [WebCite Cache ID 62zwuXUxD]
36. Fernandez-Luque L, Elahi N, Grajales FJ. An analysis of personal medical information disclosed in YouTube videos created by patients with multiple sclerosis. Stud Health Technol Inform 2009;150:292-296. [Medline: 19745316]

37. Fernandez-Luque L, Karlsen R, Melton GB. HealthTrust: trust-based retrieval of YouTube's diabetes channels. In: Proceedings. 2011 Presented at: CIKM '11: 20th ACM International Conference on Information and Knowledge Management; Oct 24-28, 2011; Glasgow, UK. [doi: 10.1145/2063576.2063854]
38. Cohen JA. Coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr;20(1):37-46. [doi: 10.1177/001316446002000104]
39. Comprehensive R Archive Network. 2012. Psych: Procedures for Psychological, Psychometric, and Personality Research URL: http://cran.r-project.org/web/packages/psych/index.html [accessed 2012-01-30] [WebCite Cache ID 655E6HDfK]
40. Buettcher S, Clarke L, Cormack G. Information Retrieval: Implementing and Evaluating Search Engines. Cambridge, MA: The MIT Press; 2010:408.
41. Järvelin K, Kekäläinen J. Cumulated gain-based evaluation of IR techniques. ACM Trans Inf Syst 2002;20(4):422-446. [doi: 10.1145/582415.582418]
42. Rodgers JL, Nicewander WA. Thirteen ways to look at the correlation coefficient. Am Stat 1988 Feb;42(1):59-66. [doi: 10.2307/2685263]
43. Davidson J, Liebald B, Liu J, Nandy P, Van Vleet T, Gargi U, et al. The YouTube video recommendation system. In: Proceedings. 2010 Presented at: Fourth ACM Conference on Recommender Systems; Sept 26-30, 2010; Barcelona, Spain. [doi: 10.1145/1864708.1864770]

Abbreviations

API: application programming interface
DCG: discounted cumulative gain
CDC: Centers for Disease Control and Prevention
HITS: hyperlink-induced topic search

Edited by G Eysenbach; submitted 06.11.11; peer-reviewed by K Denecke, A Lau; comments to author 27.11.11; revised version received 16.01.12; accepted 17.01.12; published 31.01.12.

Please cite as:
Fernandez-Luque L, Karlsen R, Melton GB
HealthTrust: A Social Network Approach for Retrieving Online Health Videos
J Med Internet Res 2012;14(1):e22
URL: http://www.jmir.org/2012/1/e22/
doi:10.2196/jmir.1985
PMID:22356723

Original Paper

De-identification Methods for Open Health Data: The Case of the Heritage Health Prize Claims Dataset

Khaled El Emam[1,2], PhD; Luk Arbuckle[1], MSc; Gunes Koru[3], PhD; Benjamin Eze[4], MSc; Lisa Gaudette[4], MSc; Emilio Neri[1], BEng; Sean Rose[4], MSc; Jeremy Howard[5]; Jonathan Gluck[6]

[1]Electronic Health Information Laboratory, CHEO Research Institute, Inc., Ottawa, ON, Canada

[2]Department of Paediatrics, University of Ottawa, Ottawa, ON, Canada

[3]Department of Information Systems, University of Maryland Baltimore County, Baltimore, MD, United States

[4]Privacy Analytics, Inc., Ottawa, ON, Canada

[5]Kaggle, Melbourne, Australia

[6]Heritage Provider Network, Northridge, CA, United States

Corresponding Author:
Khaled El Emam, PhD
Electronic Health Information Laboratory
CHEO Research Institute, Inc.
401 Smyth Road
Ottawa, ON, K1H 8L1
Canada
Phone: 1 738 4181
Fax: 1 731 1374
Email: kelemam@uottawa.ca

Abstract

Background: There are many benefits to open datasets. However, privacy concerns have hampered the widespread creation of open health data. There is a dearth of documented methods and case studies for the creation of public-use health data. We describe a new methodology for creating a longitudinal public health dataset in the context of the Heritage Health Prize (HHP). The HHP is a global data mining competition to predict, by using claims data, the number of days patients will be hospitalized in a subsequent year. The winner will be the team or individual with the most accurate model past a threshold accuracy, and will receive a US $3 million cash prize. HHP began on April 4, 2011, and ends on April 3, 2013.

Objective: To de-identify the claims data used in the HHP competition and ensure that it meets the requirements in the US Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule.

Methods: We defined a threshold risk consistent with the HIPAA Privacy Rule Safe Harbor standard for disclosing the competition dataset. Three plausible re-identification attacks that can be executed on these data were identified. For each attack the re-identification probability was evaluated. If it was deemed too high then a new de-identification algorithm was applied to reduce the risk to an acceptable level. We performed an actual evaluation of re-identification risk using simulated attacks and matching experiments to confirm the results of the de-identification and to test sensitivity to assumptions. The main metric used to evaluate re-identification risk was the probability that a record in the HHP data can be re-identified given an attempted attack.

Results: An evaluation of the de-identified dataset estimated that the probability of re-identifying an individual was .0084, below the .05 probability threshold specified for the competition. The risk was robust to violations of our initial assumptions.

Conclusions: It was possible to ensure that the probability of re-identification for a large longitudinal dataset was acceptably low when it was released for a global user community in support of an analytics competition. This is an example of, and methodology for, achieving open data principles for longitudinal health data.

(J Med Internet Res 2012;14(1):e33) doi:10.2196/jmir.2001

KEYWORDS

Open data; de-identification; privacy

Introduction

Creating open data is considered an important goal in the research community. Open data is said to ensure accountability in research by allowing others access to researchers' data and methods [1-4]. Having research data available to peers and the public helps to ensure that reported study results are valid and protects against faulty data [1-8]. Another asset of open data in research is that it allows researchers to build on the work of others more efficiently and helps to speed the progress of science [2,3,5]. To build on previous discoveries, there must be trust in the validity of prior research. Openness of research methods, as well as of raw data, facilitates trust between researchers and with the public [2]. Having research data available to other researchers allows for secondary analyses that expand the usefulness of datasets and the resulting knowledge gained [1,3-5]. Connected to this is the decrease in the burden on research participants through the reuse of existing research data and the decrease in the cost of data collection [1,3,5].

Although there is some evidence that sharing raw research data increases the citation rate of research papers [9], researchers do have concerns about open data, including the privacy of research participants, which could prevent them from sharing data [1,3,4,6,8,10-12]. Research participants put their trust in the research team to protect their privacy and keep their information confidential [3,4,8,12,13].

There is a dearth of articles documenting methods for the creation of open health data that specifically address these privacy concerns. We provide a case study of de-identifying a health dataset for public release in the context of the Heritage Health Prize (HHP).

The Heritage Health Prize

In April 2011 the Heritage Provider Network (HPN), a health maintenance organization based in California, launched the largest public health analytics competition to date: the HHP [14]. The objective of the competition is to construct a model to predict the number of days a patient will be hospitalized in the following year, by using the current and previous years' claims data. The core dataset consists of 3 years of de-identified HPN data on 113,000 patients. At the time of writing there were 1347 entrants in the competition and close to 10,000 entries. The patient data are provided to all entrants through a download on the competition website. The individual or team that develops the most accurate prediction model past a certain accuracy threshold after the 2-year competition period gets a US $3 million cash prize, or a $0.5 million prize for the most accurate model if no entrant beats the accuracy threshold.

The public disclosure of health data for the purposes of attracting data analysts from around the globe to solve complex problems or to bring rapid advances to a field is not new. Table 1 [15-18] summarizes three recent health competitions that made data publicly available. However, the privacy of patients is an important consideration when publicly disclosing a large health dataset accessible with few restrictions. In particular, there is a risk that patients in the competition dataset can be re-identified by an adversary. Re-identification can potentially harm these patients, from social and psychological harm, to financial harm by affecting their employability or insurability.

In the United States there is no legislative requirement to obtain patient consent to disclose health information if the data are deemed de-identified. The Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule provides some definitions and standards for the de-identification of health data. Therefore, a credible claim must be made that the data are indeed de-identified according to one of those standards to allow their disclosure for the HHP without obtaining patient consent.

Table 1. Recent examples of public releases of health data for the purpose of competitions.

Competition	Objective
Predict HIV Progression [15]	Finding markers in the human immunodeficiency DNA sequence that predict a change in the severity of the infection
INFORMS data mining contest [16]	Predicting hospitalization outcomes of transfer and death
Practice Fusion medical research data [17,18]	Developing an application to manage patients with a focus on chronic diseases

De-identification of the HHP Data

We describe how the HHP data were de-identified for the competition to (1) make the data publicly available, and (2) meet the requirements of the HIPAA Privacy Rule. Only one previous study explained the methods for de-identifying public-use health data files; however, it considered risks to Canadian patients and did not involve longitudinal data [19]. The main objective of the de-identification was to protect the identity of the patients.

The contributions of this work are (1) a description of how we measured re-identification risk for the public release of a large health dataset in the United States (which can be a useful example for other open government and open data initiatives and programs in the United States [20]), including the development of new risk measurement techniques, (2) an extension of an existing algorithm developed for cross-sectional data, optimal lattice anonymization, with new methods to efficiently de-identify longitudinal data, and (3) a description of an approach for using simulated attacks to evaluate longitudinal data de-identification algorithms and an illustration of its use on the HHP dataset. Our results demonstrate that it is possible to publicly disclose longitudinal health data with strong guarantees of privacy, as defined in HIPAA. Furthermore, the methods described here can serve as a template for other open competitions and the creation of open health datasets.

Methods

The competition data consist of 3 years' worth of demographic and claims data. For year 1 and year 2, the number of days of hospitalization in the subsequent year is also included. The claims data represent the predictors, and the number of days of hospitalization is the outcome. These data are used for training prediction models. Entrants use the year-3 claims data to predict the number of days of hospitalization for year 4, and the competition will be judged on the accuracy of that year-4 prediction. Therefore, entrants download the data for years 1–3, to predict days of hospitalization for year 4.

Managing the re-identification risk for the competition dataset consists of a combination of technical and legal measures. These measures are described in the following section.

Definitions

Quasi-identifiers

Quasi-identifiers are variables that represent the background knowledge about patients in the competition data that an adversary could use for re-identification. If an adversary does not have certain background knowledge, then those variables cannot be quasi-identifiers. General examples of quasi-identifiers are sex, date of birth or age, location information (such as zip codes), language spoken at home, and ethnic origin.

Equivalence Classes

All records that share the same quasi-identifier values are called an *equivalence class*. For example, all the records in a dataset about 17-year-old males admitted on January 1, 2008 are an equivalence class. Equivalence class sizes for a quasi-identifier (such as age) could potentially change during de-identification. For example, there may be 3 records for 17-year-old males admitted on January 1, 2008. When the age is recoded to a 5-year interval, then there may be 8 records for males between 16 and 20 years old admitted on January 1, 2008. In general there is a trade-off between the level of detail provided for a quasi-identifier and the size of the corresponding equivalence classes, with more detail being associated with smaller equivalence classes.

Identity Versus Attribute Disclosure

Two kinds of disclosure are of concern. The first occurs when an adversary can assign an identity to a record in the disclosed dataset. For example, if the adversary is able to determine that record number 7 belongs to patient Alice Smith, then this is called *identity disclosure*. The second type of disclosure happens when an adversary learns something new about a patient in the database without knowing which specific record belongs to that patient. For example, if all 20-year-old female patients in the disclosed database who live in a particular county have a particular diagnosis, then an adversary does not need to know which record belongs to 20-year-old Alice Smith, if she lives in that county, to know that she has that particular diagnosis. This is called *attribute disclosure*.

All known re-identification attacks of personal information that have actually occurred have been identity disclosures [21]. Furthermore, the HIPAA Privacy Rule is concerned only with protecting against identity disclosure. Consequently, identity disclosure was the primary risk that needed to be addressed. We therefore focused solely on identity disclosure for the purpose of de-identification.

Dataset

The claims dataset consists of two tables that include the fields shown in Table 2 and Table 3 [22,23]. The records in both tables are linked through the MemberID field. The *patients* table has only 1 record per patient. The *claims* table contains records for all of the patient claims included in the dataset. Patients have different numbers of claims over the 3-year period.

The quasi-identifiers included in the dataset (indicated in Table 2 and Table 3) are the variables that we believed could be used by an adversary for a re-identification attack. The justification for this selection of quasi-identifiers will become evident below when we discuss the plausible re-identification attacks that could be made on these data.

Table 2. Description of the fields in the patients data table.

Field	Description
MemberID	Unique identifier for the patient
Age[a]	Age in years at the time of the first claim in year 1
Sex[a]	Patient's sex
DaysInHospital Y2[a]	Total number of days the patient was hospitalized in year 2
DaysInHospital Y3[a]	Total number of days the patient was hospitalized in year 3

[a] Quasi-identifier.

Table 3. Description of the fields for the claims data table.

Field	Description
MemberID	Unique identifier for the patient
ProviderID	Unique identifier for the responsible provider giving care
Vendor	Unique identifier for the vendor providing the service
PCP	Unique identifier for the primary care provider
Year	Indicator of claim year (year 1, year 2, or year 3)
Specialty[a]	Specialty of provider
PlaceOfService[a]	Place of service
CPTCode[a]	CPT[b] code: these codes provide a means to accurately describe medical, surgical, and diagnostic services, are used for processing claims and for medical review, and are the national coding standard under HIPAA[c]
LOS[a]	Length of stay in hospital
DSFC[a]	Number of days since first claim computed from the first claim for that patient for each year
PayDelay	Number of days of delay between date of service and date of payment of the claim
Diagnosis[a]	ICD-9-CM[d] code

[a] Quasi-identifier.

[b] Current Procedural Terminology [22].

[c] Health Insurance Portability and Accountability Act.

[d] *International Classification of Diseases*, 9th revision, Clinical Modification [23]

Preprocessing of Claims Data

We preprocessed the data to apply some basic de-identification steps before assessing any quantitative re-identification risk.

Creating Pseudonyms

The MemberID, ProviderID, Vendor, and PCP fields were converted to irreversible pseudonyms [24], since they would otherwise be considered direct identifiers. These ID values are used during the provision of care and therefore are generally known. Consequently, these direct identifiers could potentially be exploited for financial gain, and were therefore pseudonymized. For example, the original IDs could be used to identify individual providers and the number and type of procedures that they perform.

Top-Coding

Quantitative values that are considered uncommonly high are often limited to an upper bound, a procedure called *top-coding*. Such extreme values make individuals more unusual (or make them stand out) in the population and can be revealing by themselves or used to infer other characteristics about the patients that should be protected.

A commonly used heuristic for top-coding is to have a cut-off at the 99.5th percentile [25]. To err on the conservative side, we top-coded the PayDelay variable and the DaysInHospital variable at the 99th percentile. Extreme values on PayDelay could indicate procedures that are more expensive and for which it would take an unusually long time to pay (and hence the procedure could be inferred from the PayDelay value). Extreme values for DaysInHospital could cause patients to stand out because they have stayed exceedingly long in hospital.

Truncation of Claims

While it is not likely that an adversary would know the exact number of claims that an individual patient would have, it is plausible for an adversary to know whether an individual patient has had an abnormally large number of claims. For example, a patient may have 300 claims a year and be the only one in the population with more than 200 claims. Adversaries who know that their 50-year-old neighbor has had an unusually high number of hospital procedures could correctly guess that this extreme outlier is their neighbor.

We therefore truncated the number of claims per patient at the 95th percentile. To decide which claims to truncate we assigned each claim a score, and deleted claims with the highest scores from the dataset. A description of the scoring method is provided in Multimedia Appendix 1.

The truncation of claims was different from the censoring method that has been described in previous research for diagnosis codes [26]. The censoring method collapses repeating codes and then uses suppression to ensure a minimal number of patients have the same code. In our case we did not collapse similar claims, and our focus was on full claims that contained multiple pieces of information (as summarized in Table 3), as opposed to just diagnosis codes.

Removal of High-Risk Patients and Claims

Patients who were considered to be high risk were removed from the dataset to avoid the chance that their disease, condition, or procedure could be inferred from patterns in the data. These patients had *International Classification of Diseases*, 9th revision, clinical modification (ICD-9-CM) or Current Procedural Terminology (CPT) codes that represent highly

stigmatized conditions, and conditions that patients who lock their records (through consent directives) tend to have and want to conceal—for example, patients with human immunodeficiency virus infection, those who have had abortions, and patients with rare and visible diseases and conditions [27]. For those individuals HPN deemed that the only acceptable risk of re-identification was zero. The criteria used to remove patients are listed in Multimedia Appendix 1.

As Multimedia Appendix 1 shows, the patients who were removed during this step were different on several factors from the rest of the patient population: (1) their length of stay (LOS) in hospital tended to be longer, (2) they tended to be older, (3) the interval between their claims was longer, and (4) they tended to have more claims. This suggested that they had more serious chronic conditions than the rest of the population dataset. On the other factors there was little or no difference.

Suppression of Provider, Vendor, and PCP Identifiers

Providers could have patterns of treatment that make them stand out. An adversarial analysis by an independent party of a prerelease version of the HHP dataset noted how information about providers could potentially be used to predict the hospitals where procedures were performed (A Narayanan, unpublished data, 2011). Knowledge of the treating hospital would increase the risk of re-identification for the patients.

These patterns of treatment consisted of 4 quasi-identifiers: the place of service, specialty, CPT code, and diagnosis code. For example, a provider could be the only one with a particular specialty in a specific place of service who performed procedures on patients with a particular diagnosis. In cases where it was estimated from the HHP data that there were fewer than 20 providers with the same pattern in the HPN system, the provider ID was suppressed for those records. The choice of 20 is justified below in the section outlining thresholds. The estimation method used is described elsewhere [28,29]. A similar process was followed for vendor and primary care provider IDs.

Facts and Assumptions About Re-identification Threats

To understand the type of de-identification required to protect patients, we first had to determine the threats that could exist for the duration of the competition. The following are the key facts and assumptions of the threat modeling used:

- Fact: The dataset that was being released for the HHP consisted of a small sample of all HPN patients.
- Fact: All entrants in the competition had to sign (or click through) an agreement saying that they would not attempt to re-identify patients in the dataset, contact patients, or link the HHP data with other datasets that would add demographic, socioeconomic, or clinical data about the patients (where such data could make the risk of re-identification much higher).
- Assumption: It would not be possible for an adversary to know whether the record for a particular patient was in the HHP dataset. If an adversary made a guess, it would be equal to the sampling fraction. Most patients would themselves not know whether they were members of HPN, and therefore the most realistic sampling fraction to use would be from the population of counties in California covered by HPN. However, to err on the conservative side, we assumed that an adversary would know whether a patient was a member of HPN in our calculations of re-identification risk.
- Assumption: An adversary would have background information about only a subset of the claims of a patient in the dataset. For example, if a patient had 100 claims, we did not deem it plausible for the adversary to know the exact information in all of those 100 claims and to use that information for re-identification purposes. Rather, we assumed the adversary would have information about only a subset of these claims. This has previously been referred to as the *power* of the adversary, and various methods have been used to account for power when de-identifying transactional data [30-32].

These facts and assumptions shaped how we conceptualized re-identification risk and which kinds of attacks we considered plausible for this dataset.

Attacks on the Data

We examined plausible attacks on the data as described below, and for each one we will discuss how we measured and managed the re-identification risks.

One important distinction to make at the outset pertains to subcontractors (eg, insurers, laboratories, or pharmacists) and employees of HPN, versus the entrants. Subcontractors process patient data during the regular provision of care and will have a large amount of information about the patients in the competition that can potentially be used for re-identification. However, HPN has contracts with these subcontractors and there are already mechanisms in place to enforce these agreements. In such a case, reliance on existing legal methods to protect against re-identification by subcontractors was deemed sufficient.

On the other hand, entrants in the competition could come from many countries in the world. Even though entrants had to agree to a certain set of rules, enforcement of the rules globally poses a practical challenge.

Therefore, we assumed that an adversary would be one of the entrants who has obtained the HHP data (1) by registering for the competition, or (2) through a data leak (deliberate, accidental, or malicious) from a legitimate entrant. Furthermore, it would not be prudent to assume that the adversary would adhere to conditions on other public or semipublic databases to which they have gained access. In such a case, we needed technical methods that provide stronger guarantees that the probability of re-identification is low.

Attack 1: The Nosey Neighbor Adversary

Under this attack, the adversary would be an individual who (1) would be trying to re-identify a target individual who was an HPN patient (a specific individual, such as a neighbor or a famous person) or any individual who was known to the adversary to be an HPN patient (an arbitrary individual selected at random), (2) would not know whether the target individual

was in the dataset, and (3) would have some basic background information about the target patient in terms of the patient's demographics and information about some of the patient's claims.

The adversary could be a patient's neighbor, coworker, relative, or ex-spouse, or the target individual could be a famous person whose basic demographics and perhaps some of whose treatment information would be publicly known. There are known examples of this kind of attack. In one case a researcher re-identified the insurance claim transactions of the Governor of Massachusetts [33]. In another example, a neighbor was re-identified in a hospital prescription database that was going to be disclosed to a commercial data broker [34].

Under this type of attack, the risk metric would be the probability that an individual can be correctly re-identified. The probability of an individual being re-identified using this attack is the reciprocal of the equivalence class size in the HPN member population (from which the competition dataset is derived) [28].

For any patient in an equivalence class j, the probability of re-identification was defined as equation 1 (Figure 1), where F_j is the equivalence class size in the HPN patient population. Since we did not know which record might be attacked, we used the record with the highest risk as a risk measure for the whole file (equation 2, Figure 1).

Figure 1. Equations describing how re-identification risk was measured.

$$1/F_j \quad (1)$$

$$1/\min_j(F_j) \quad (2)$$

$$k = \alpha F_j' \quad (3)$$

$$100\times\frac{1}{N}\sum_{j\in J} f_j \times I(F_j \geq 20) \leq 0.8 \quad (4)$$

$$100\times\frac{1}{N}\sum_{j\in J} f_j \times I(f_j \geq 20\alpha) \leq 0.8 \quad (5)$$

Attack 2: Matching With the Voter Registration List

In California it is possible to obtain the voter registration list [35]. The voter registration list contains the date of birth and gender of the voter. We did not include the full date of birth in the HHP dataset, but we did include a generalization of the date of birth to a 10-year interval. Even though there are restrictions on what a California voter registration list can be used for [35], an adversary could potentially match the HHP dataset with the voter list to re-identify patients. In such a case, the appropriate re-identification risk metric would be the proportion of individuals that could be re-identified in the HHP dataset (this metric is also known as marketer risk [29]).

It has been shown that managing the risk in equation 2 also manages marketer risk [29]. And, because the equivalence classes in the voter registration list are the same size as or larger than the equivalence classes in the HPN population, if we could protect against attack 1, then we would automatically protect against attack 2.

Attack 3: Matching With the State Inpatient Database

In the United States, 48 states collect data on inpatients [36], and 26 states make their hospital discharge data available through the Agency for Healthcare Research and Quality (AHRQ) [37]. These data can be purchased for the purposes of research or another approved use. These datasets are referred to as the State Inpatient Database (SID).

An adversary could potentially match the competition dataset with the SID data to discover something new about the individuals in the dataset. For example, if an individual were able to match the HHP records with the SID records, then the adversary could discover the exact month and year of birth of patients and their detailed diagnosis codes and procedures, even if we generalized them in the HHP data release (since these fields are included in the SID). Furthermore, the SID contains race information, which could be added to the HHP dataset after matching. This would provide more detailed information than was disclosed in the HHP dataset and would therefore raise the re-identification risk for any correctly matched patients.

Note that not all patients in the HHP dataset were hospitalized. Some may, for example, have been seen in an outpatient clinic. Therefore, by definition only a subset of the HHP dataset could be matched with the SID.

For this attack, the re-identification risk metric would be the proportion of individuals that could be matched between the HHP and the SID datasets. This can be measured using the marketer risk metric [31].

Since the SID covers all hospital discharges in California, the equivalence class sizes for hospitalized patients in the HPN population were equal to or smaller than the SID equivalence classes for those patients. This means that if we managed the risk in equation 2 for attack 1, we would also manage the risk for attack 3.

Summary of Re-identification Risks from Attacks

Based on the above analysis of the various possible attacks, if the re-identification risk from attack 1 could be managed, then the risks from all of the other attacks would also be managed. Below we describe the algorithm used in this study to manage the risk from attack 1. Additionally, during the empirical evaluation component of our study, we measured the re-identification risks from attacks 1 to 3 to confirm that the re-identification risks for all three attacks were acceptably low.

Methods for the De-identification of the HHP Dataset

We used an automated algorithm to de-identify the dataset through generalization. Our base automated de-identification algorithm was OLA [38]. OLA provides a globally optimal solution and has been shown to have good performance on real health datasets [38]. It is a k-anonymity algorithm. The

k-anonymity criterion is one of the most common ways to de-identify a dataset [38-42] and can be used to manage the probability of re-identification due to identity disclosure [28]. OLA has been designed to work only on cross-sectional data. As we describe further below, we have extended this algorithm to de-identify longitudinal data. We refer to our extended algorithm as longitudinal OLA (LOLA).

Base Algorithm

We will provide a brief overview of how LOLA works and its parameters, and then explain how we modified these parameters for the de-identification of the longitudinal HHP dataset.

Input

LOLA has two inputs. The first is the *k* value, which indicates the maximum amount of re-identification risk the data custodian is willing to take, and this determines the amount of de-identification that will be applied. The *k* value is the minimum size of an equivalence class. This means that the maximum probability of a record being correctly re-identified is given by 1/*k* (this is the risk threshold). LOLA's second input pertains to the percentage of records that have a risk higher than the risk threshold: the *MaxSup* parameter.

In our case we defined $k = f_j$, where the f_j value is the minimum equivalence class size in the HHP dataset. Since the risk we wanted to manage was $1/F_j$, which was based on the equivalence class sizes in the HPN member population, we made the large sample assumption that $f_j = \alpha F_j$, where α is the sampling fraction and F_j is the smallest equivalence class size in the HPN population. Therefore, we defined equation 3 (Figure 1).

For example, if we had set $F_j = 20$ and a 20% sampling fraction was used, then the *k* value for LOLA would have been 4.

Note that in practice more sophisticated methods for estimating f_j would be used as described elsewhere [28,29], especially for small sampling fractions. In our case, we used an estimator based on the truncated Poisson distribution [28].

Generalization

A key step in LOLA is generalization. Generalization reduces the precision in the data. As a simple example, a patient's date of birth can be generalized to the month and year of birth, to the year of birth, or to a 5-year interval. Allowable generalizations are specified in generalization hierarchies. Let us consider an example dataset with only 3 quasi-identifiers: date of birth (d), gender (g), and date of visit (p). Figure 2 shows the domain generalization hierarchies for these quasi-identifiers. These hierarchies describe how the precision of each quasi-identifier can be reduced during generalization.

All of the possible generalizations can be expressed in the form of a lattice as shown in Figure 3. In this lattice each possible generalization is represented by a node starting from the original dataset at the lowest node, $<d_1,g_1,p_1>$. As one moves up the lattice, the quasi-identifiers are generalized. For example, node $<d_2,g_1,p_1>$ has the date of birth generalized to month and year. The objective of LOLA is to efficiently find the best generalization solution (node) in that lattice. The best node meets two criteria: (1) the proportion of records that are considered to have a high probability of re-identification is less than or equal to *MaxSup*, and (2) the best node has the smallest amount of information loss.

After efficiently evaluating the nodes in the lattice, LOLA identifies the candidate nodes that meet criterion 1 above. Out of the candidate nodes, LOLA then chooses the node with the smallest information loss among the candidate nodes, and this meets criterion 2. Information loss is measured in terms of a general entropy metric, which was found to have properties superior to those of other commonly used metrics in the literature [38].

Figure 2. The three domain generalization hierarchies for the 3 quasi-identifiers: date of birth (d), gender (g), and visit date (p).

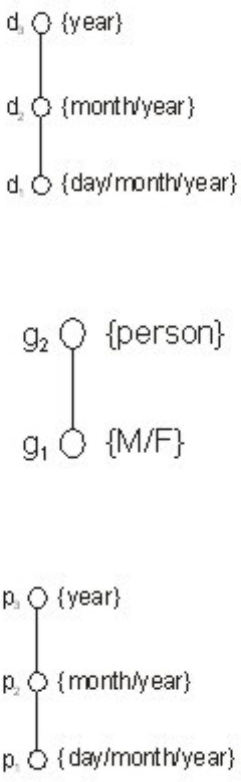
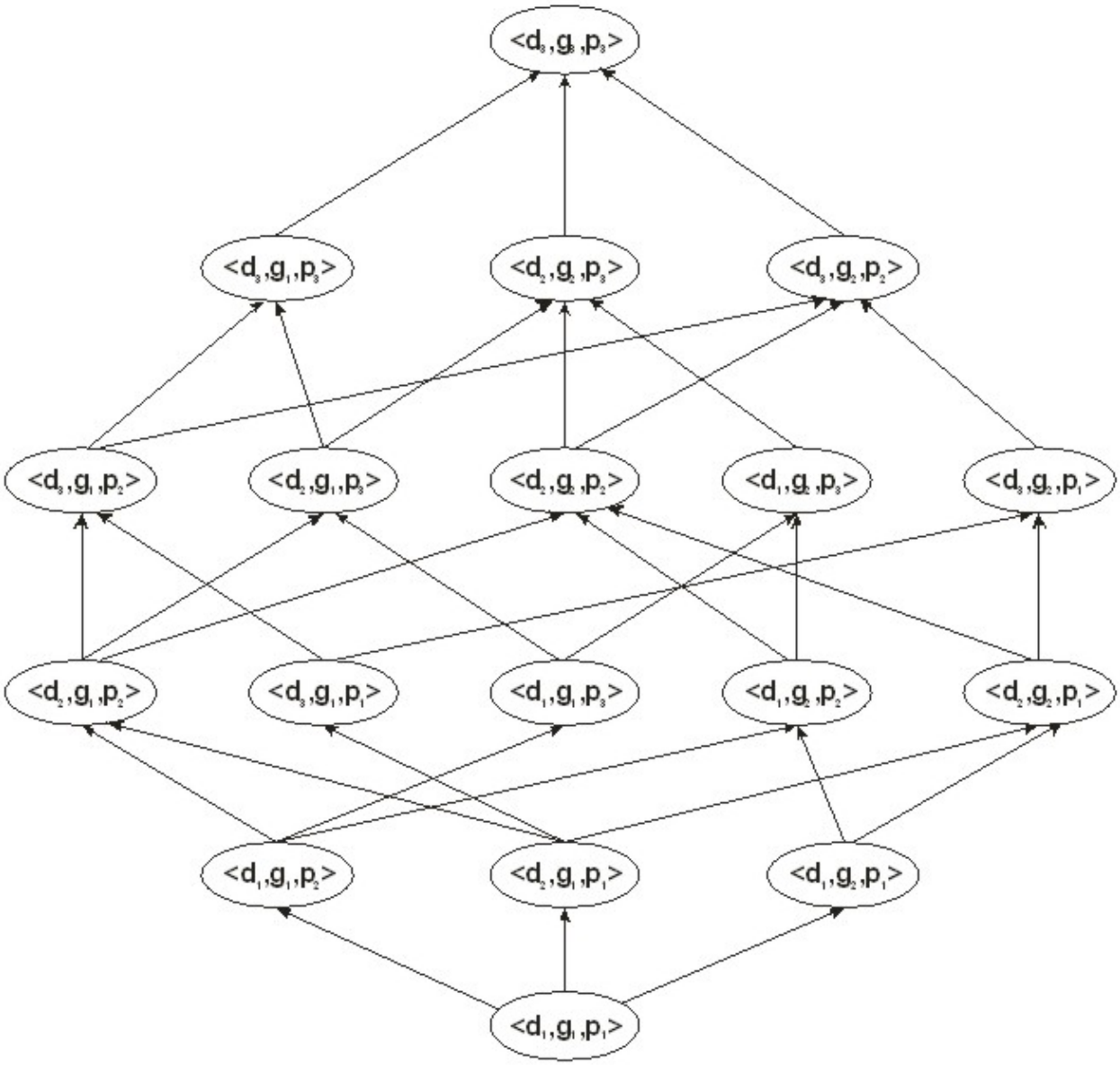


Figure 3. A lattice showing the possible generalizations of the 3 quasi-identifiers: date of birth (d), gender (g), and visit date (p).



Acceptable Re-identification Risk

In the United States, the HIPAA Privacy Rule Safe Harbor de-identification standard was conceptualized using population uniqueness of individuals as the measure of risk, as documented in the responses to comments by the Department of Health and Human Services [43,44]. The population uniqueness standard means that individuals who are the only one in their equivalence class are at high risk, but if there is more than one person in the equivalence class then they are not considered at a high risk of re-identification. The designers of Safe Harbor considered very small" risk in terms of the percentage of the population that is unique. Studies have shown that in datasets that meet the Safe Harbor de-identification standard, 0.04% of the population is unique [45,46]. HPN wanted to ensure that the risk exposure for the HHP dataset was equal to or less than the Safe Harbor risk exposure. Risk exposure is defined as *loss* × *probability*.

Risk exposure only comes from the records that have an unacceptably high probability of re-identification. In our case, the *loss* pertained to the number of individuals who could be re-identified (ie, individuals at a high risk of re-identification), and *probability* meant the probability of re-identification. Therefore, for a dataset with N records, the maximum risk exposure under Safe Harbor would be $0.004 \times N \times 1$.

A uniqueness threshold would be considered quite high by most standards (for example, see [13,47-60]). HPN decided that a probability of re-identification for *a record* in the HHP dataset of .05 was an acceptable risk. The choice of this value was informed by two precedents. The Center for Medicaid & Medicare Services has created public-use files containing claims data from a sample of beneficiaries while ensuring that the probability of an individual being re-identified was less than .1 [61]. An analysis to prepare a public discharge abstract dataset from Canadian hospitals used a threshold probability of .05 [19]. The value chosen by HPN erred on the more conservative side of this range.

This means that, if the probability was equal to or lower than .05, then the data would be acceptable for release. To ensure a risk level that low we needed to ensure that $F_j \geq 20$ for all j (ie, $F_j = 20$ in equation 3).

If we revisit our definition of k in equation 3, we would have $k = 30\alpha$, since this threshold translates to a maximum equivalence class size of 20 in the population.

To retain the same level of maximum risk exposure as Safe Harbor with our proposed .05 probability threshold, we could accept only 0.8% of the records to have a probability that was higher than the .05 threshold for the same value of N (ie, $.05 \times 0.008 \times N = 1 \times 0.004 \times N$). The 0.8% value then represented the *MaxSup* that was used in the LOLA algorithm.

Therefore, if the condition in equation 4 (Figure 1) was met, then we considered the risk acceptable. In equation 4, $I(\cdot)$ was the indicator function. Making the large sample simplification resulted in equation 5 (Figure 1), although, as noted above, better estimates of F_j could be used for small samples [28,29].

Generalization Hierarchies for the HHP Dataset

As Table 4 shows, 4 quasi-identifiers had generalization hierarchies: Age, DaysInHospital, LOS, and DSFC. These were the ones we used to construct the lattice for the application of LOLA. For the remaining quasi-identifiers there was only one level of predetermined generalization.

Each claim had up to 4 diagnosis codes. These were converted into 2 values. The ICD-9-CM diagnosis codes were generalized into 45 primary condition groups, which have been determined to be good predictors of mortality [62]. We also created a categorized comorbidity score (Charlson index) [63,64]. The CPT codes were generalized to a higher code in the CPT hierarchy. In consultation with clinical experts, we also grouped values within the Specialty and PlaceOfService variables. All groupings are described in Multimedia Appendix 1. Table 4 summarizes all of the generalization hierarchies used.

Table 4. Description of the generalization hierarchies for the quasi-identifiers.

Quasi-identifier	Description
Age	Years → 5-year interval; 80+ → 10-year interval; 80+ → 20-year interval; 80+
Sex	no change
DaysInHospital Y2/Y3	Days → days to 2 weeks; >2 weeks → days to 1 week; 1–2 weeks; >2 weeks
Specialty	Original specialty → grouped specialty (see Multimedia Appendix 1)
PlaceOfService	Original place of service → grouped place of service (see Multimedia Appendix 1)
CPTCode[a]	Original CPT code → grouped CPT code
LOS[b]	Days → days up to 6 days, weeks afterward → days up to 6 days; (1–2] weeks; (2–4] weeks; (4–8] weeks; (8–12 weeks]; (12–26] weeks; 26+ weeks → <1 week; (1–2] weeks; (2–4] weeks; (4–8] weeks; (8–12 weeks]; (12–26] weeks; 26+ weeks → <4 weeks; (4–8] weeks; (8–12 weeks]; (12–26] weeks; 26+ weeks
DSFC[c]	Days → weeks → 2 weeks → months
Diagnosis	ICD-9-CM[d] code → primary condition group (see Multimedia Appendix 1)

[a] Current Procedural Terminology.

[b] Length of stay in hospital.

[c] Days since first claim.

[d] *International Classification of Diseases*, 9th revision, Clinical Modification.

Adversary Knowledge and Power

The *power* of the adversary reflects the number of claims that the adversary would have background information about, and it pertains to the claims data and not to the basic information about the patients (such as their demographics). We denoted the power of the adversary as p. If each claim had only 1 quasi-identifier, then p would mean that the adversary had knowledge about the value for that quasi-identifier in p claims. With more than 1 quasi-identifier in each claim, the adversary would know each of the quasi-identifier values for p claims.

Previous research that considered the power of the adversary always assumed that the power is fixed for all patients [30-32,65,66]. However, intuitively it makes sense that the adversary would have different amounts of background knowledge, or power, for different patients. For example, everything else being equal, it is easier to have background information about a patient with a large number of claims than about a patient with few claims. Therefore, we would expect power to increase monotonically with the number of claims that a patient has.

Also, it is likely that certain pieces of background information are more easily knowable than others by an adversary, making it necessary to treat the quasi-identifiers separately when it comes to computing the power of an adversary. For example, it would be easier to know a diagnosis value for patients with chronic conditions whose diagnoses keep repeating across claims. In such a case, if the adversary knew the information in 1 claim, then it would be easier to predict the information in other claims, increasing the amount of background knowledge that the adversary can have. In this case the diversity of values on a quasi-identifier across a patient's claims becomes an

important consideration. Therefore, we expect the power of an adversary to decrease monotonically with the diversity of values on the quasi-identifiers.

As Multimedia Appendix 1 shows, we found that the correlations among the diversity of quasi-identifiers tend to be small to moderate. For example, the correlation between the diversity of Specialty and LOS was .08. This means that there is a very weak relation between the diversity in these two variables. On the other hand, the correlation between the diversity of Specialty and CPTCode was .4 and that between LOS and DSFC was –.44. In the former case it means that the more varied the specialty of the treating physician, the more varied the procedures. In the latter case, the more varied the LOS, the less varied the time between claims. These relationships make sense, but they also suggest that it would be more appropriate to compute a power for each quasi-identifier separately, because diversity is not uniform across quasi-identifiers for the same patient.

We defined the power for a particular individual in the data and for a particular quasi-identifier as p_{ih}, where i represents the individual and h represents the quasi-identifier. In Multimedia Appendix 1 we present a method for computing a value for p_{ih} that takes into account the number of claims and diversity. We also set the maximum power p_m as $\max(p_{ih}) = 5$. This means that, for any single patient, the maximum number of values (claims) that an adversary would have is 5 for each quasi-identifier. For instance, if there are 2 quasi-identifiers, then the adversary can have a maximum of 10 pieces of information on the patient. In our empirical evaluation we assessed the sensitivity of our results to this value.

We made two assumptions about the knowledge of the adversary: (1) the adversary would not know which values on the quasi-identifiers were in the same claim (the *inexact knowledge* assumption), and (2) the adversary would not know the exact order of the claims (the *inexact order* assumption) beyond what is revealed through the DSFC quasi-identifier, which is consistent with other models of transactional data in the disclosure control literature [30-32,65,66]. However, we did test the sensitivity of our results to these assumptions in our empirical evaluation.

Node Computation

As noted earlier, the LOLA algorithm performs an efficient search through the lattice. During this search it needs to evaluate the percentage of records that are high risk for some of the nodes in the lattice. This is called *node evaluation*.

It would have been computationally very expensive for us to evaluate all combinations of p_{ih} values for each quasi-identifier. For example, computing all combinations of 5 values from, say, 100 claims would have required more than 75 million computations of risk. Since these 5 combinations would be different for each patient, this would need to be repeated tens of thousands of times.

Therefore, we used a hierarchical bootstrapping approach [67]. Here we sampled 10,000 patients with replacement, and for each sampled record we selected p_{ih} quasi-identifier values across the claims without replacement. We then computed the proportion of patients who were at high risk in each iteration and took the mean across 1000 iterations. If the mean number of patients who were flagged as high risk was greater than *MaxSup*, then we did not consider the node to be a candidate solution (see equation 5, Figure 1). Multimedia Appendix 1 provides a complete description of the node computation.

Empirical Evaluation

After the de-identification of the dataset using LOLA, we wanted to empirically evaluate whether the risks from the three plausible attacks were appropriately managed. Hence, we performed an empirical evaluation.

Attack 1

To evaluate the actual probability of re-identification under this attack, we developed a separate attack program that would simulate exactly what an adversary would do. This program was developed by an independent programmer not involved in the development and application of LOLA described above.

The simulated attack assumed that an adversary would choose a patient from the HPN population at random. The adversary would not know whether that individual was in the HHP dataset, and hence this would introduce some uncertainty. If the individual was in the dataset then we computed the appropriate p_{ih} value for each quasi-identifier for that individual, and then selected the items of background knowledge about the individual. We then attempted to match the background knowledge with all of the patients in the competition dataset. The simulation was run 10,000 times, and the average match success rate gave us an estimate of the probability that an HPN patient could be correctly re-identified from the competition dataset under the assumptions that we made.

The purpose of the simulated attack was to mimic what an adversary would do. We assumed that the adversary had background information about Alice. Alice may be the adversary's neighbor or a famous person. She could also be someone the adversary selected at random from all HPN members.

The simulation dataset had two levels. Level 1 was the basic patient demographics as in Table 2. At level 2 were the quasi-identifiers in each claim, and a patient may have had a large number of claims. The level 2 data had some quasi-identifiers as shown in Table 3.

We also needed to create two versions of the de-identified dataset. Version D1 of the dataset had all of the claims for each patient. Version D2 of the dataset was the one with truncated claims. It is version D2 of the dataset that was released for the competition, but we needed D1 for the simulation. The level of generalization in the two datasets was exactly the same, the only difference being in the truncation of claims.

The following process was repeated 10,000 times:

- We drew a sample from a binomial distribution with a probability of α. This reflects the probability that an individual that the adversary knew about was in the dataset. If the value drawn was 1, then we could continue; otherwise,

we would go to the next iteration (and the current iteration was considered a failed match).
- Then we chose a target individual from the D1 dataset at random.
- We chose at random p_{ih} values for each level 2 quasi-identifier from the D1 dataset for that target individual. These values and the level 1 quasi-identifiers were the background information that the adversary would have.
- We matched that background information to the records in D2. This produced a matching equivalence class.
- One of the records was selected in the matching equivalence class at random.
- If the selected record was the correct patient then that was a successful match; otherwise, it was considered a failure.

Across the 10,000 iterations we computed the proportion of times that a correct match was found. This was the re-identification probability for the dataset taking into account the uncertainty due to the fact that we had a sample and due to the adversary not knowing which claims were truncated.

Attack 2

To compute marketer risk [29] for attack 2 we assumed that the voter registration list captured the full population, an assumption used in previous research as well [35]. We first needed to calculate the size of the equivalence classes in the population of California in the counties serviced by HPN. In this case, the equivalence classes were defined by the number of people born of each sex, in one of the counties of interest, and of each age. To compute this size, we took age and sex values from the 2000 census (the 2010 census data were not available at the time we performed this analysis), which were available at the county level for 5-year intervals (top-coded to 90+) on the American FactFinder website from the Census Bureau. This produced 20 equivalence classes per county. We derive in Multimedia Appendix 1 a closed-form equation for the expected number of records that would be correctly matched when a sample of a given size is drawn from a population. The derivation allowed us to compute expected marketer risk without having to perform an actual matching experiment or a Monte Carlo simulation.

Attack 3

We estimated the proportion of HHP records that could be correctly matched with the SID on the quasi-identifiers using the closed-form marketer risk calculation described in Multimedia Appendix 1. Here we assumed that the HHP would be a sample from the SID. The marketer risk calculations were performed for different combinations of quasi-identifiers. We assumed that the HHP dataset had all of the visits in the SID, and therefore all of the visits could be used for matching. We purchased the SID for the state of California from AHRQ for the 3 years covered by the HHP dataset to perform this analysis.

Sensitivity Analysis

We also analyzed sensitivity for the assumptions we made under attack 1. We explored three relaxations to the assumptions:

- The maximum power of the adversary, p_m, was higher than our assumed value of 5. We set the power to 10 and then to 15. With 6 quasi-identifiers, this would mean that the adversary knew up to 60 to 90 pieces of information about the patients they were attempting to re-identify.
- For 1 claim the adversary knew all of the quasi-identifiers for that claim. For example, say that we had only 2 quasi-identifiers, LOS and Diagnosis. Then we would assume that the adversary knew the LOS and Diagnosis values for the same claim. This relaxes the inexact knowledge assumption.
- The adversary knew the order of 1 pair of quasi-identifier values. For example, the adversary would know that diagnosis A preceded diagnosis B. This would apply only in cases where the power for the quasi-identifier was greater than 1. We would apply this for a pair of claims for each quasi-identifier. This relaxes the inexact order assumption.

With these three types of sensitivity analyses we believed we covered plausible scenarios in which the adversary would have extensive knowledge about the individuals in the competition dataset.

Results

The final claims dataset consisted of information from 113,000 patients, with 2,668,990 claims. The median number of claims per person was 11 and the maximum 136. Only 9556 patients had some of their claims truncated during the de-identification.

Making the conservative assumption that 0.8% of the individuals with a probability of re-identification higher than our threshold of .05 would have a probability of re-identification of 1, we would expect at most 5.8% of the patients to be re-identified based on our de-identification parameters.

After applying the LOLA algorithm to determine the optimal generalizations, we obtained the final results presented in Table 5. With these generalizations, 0.84% of the patients could be correctly re-identified using our simulated attack 1.

The risk calculation for attack 2 was that an expected proportion of 0.0005% of the HHP dataset could be correctly re-identified by matching with the appropriate counties in the California voter registration list. Furthermore, there are restrictions on the use of the California voter registration list that would prohibit such re-identification attempts [35]. Therefore, attack 2 was deemed to be very low risk.

The results for attack 3 are shown in Table 6 for various combinations of quasi-identifiers by year and across all years. As shown, the match success proportion was quite low, making the risk of gaining correct additional information about the HHP patients acceptable given our thresholds.

Table 7 shows the results of the simulation attack to evaluate sensitivity to violations of our assumptions for attack 1. The re-identification probability was not affected much by the increase in the power of the adversary. A primary reason was that many patients had 5 claims or fewer. Therefore, increasing the power did not necessarily mean that the adversary would have more background information about them. If we assume that the adversary would know which pieces of information were in the same claim, this would increase the risk, but even

at a power of 15 the probability was below what would be considered acceptable.

Table 5. Final generalizations in the dataset.

Quasi-identifier	Generalization
Age	10-year interval; 80+
Sex	No change
DaysInHospital Y2	Days to 2 weeks; >2 weeks in year 2
DaysInHospital Y3	Days to 2 weeks; >2 weeks in year 3
Specialty	Grouped specialty (see Multimedia Appendix 1)
PlaceOfService	Grouped place of service (see Multimedia Appendix 1)
CPTCode[a]	Grouped CPT code (see Multimedia Appendix 1)
LOS[b]	Days up to 6 days; (1–2] weeks; (2–4] weeks; (4–8] weeks; (8–12 weeks]; (12–26] weeks; 26+ weeks
DSFC[c]	4 weeks
Diagnosis	Primary condition group (see Multimedia Appendix 1)

[a] Current Procedural Terminology.

[b] Length of stay in hospital.

[c] Days since first claim.

Table 6. Estimated proportion of all records in the Heritage Health Prize dataset that would be correctly matched against the State Inpatient Database.

Age	LOS[a]	Sex	Number of visits	PCG[b]	CPT[c]	Year 1	Year 2	Year 3	All years
X	X	X	X			0.001612	0.001478	0.001515	0.005141
X	X	X			X	0.007105	0.005684	0.005965	0.009735
X	X	X		X		0.013334	0.010156	0.010928	0.013579
X	X	X		X	X	0.017272	0.012702	0.013797	0.015991

[a] Length of stay in hospital.

[b] Primary Condition Group.

[c] Current Procedural Terminology.

Table 7. Percentage of total records correctly matched under simulated attack with different assumptions about the number of claims (power).

	Power of adversary		
Assumption	5	10	15
Original adversary assumptions	0.84%	0.94%	1.17%
Multiple quasi-identifiers in the same claim	3.67%	3.72%	3.87%
Ordered claims	0.96%	1.0%	1.2%

Discussion

Summary

The detailed re-identification risk assessment on the HHP dataset allowed the disclosure of comprehensive longitudinal claims information on a large number of individuals while being able to make strong statements about the ability to re-identify these individuals. The de-identification we performed ensured that the risk was acceptable under different types of attacks, even to the extent that we allowed for some of our initial assumptions to be incorrect. In particular, we were able to ensure that the risk exposure was at or below the current risk exposure under the HIPAA Safe Harbor de-identification standard.

Ensuring the utility of the dataset is an important requirement in any de-identification effort. If no team is able to meet the prediction performance threshold to win the grand prize, then this may be because the threshold was too ambitious or because the de-identification itself made achieving that threshold difficult. An evaluation of the accuracy of the models before and after de-identification would be a useful exercise to help inform future competitions and fine-tune de-identification methods.

As our literature review in Multimedia Appendix 1 illustrates, existing de-identification methods for longitudinal data would not have created a dataset suitable for this competition. In that

regard, the approach presented here is one of the few available for creating public health datasets.

Limitations

Alternative ways for grouping the diagnosis and procedure codes could have been used. For example, we could have clustered the codes based on the average number of days of hospitalization. This would potentially have retained some important relationships in the data. Furthermore, it would ideally be necessary to perform this clustering using all of the quasi-identifiers to ensure that the multivariate relationships are retained. The practical challenge with such an approach was that many patients had zero days in hospital (for example, they were outpatients). This would then have resulted in coarser groupings than those we included with our analysis. Appropriate grouping of such nominal variables is an important area of future research to address constraints imposed by real datasets.

We did not consider the real possibility that there were errors in the background knowledge of the adversary. If errors exist then the match percentages would be lower than those we presented in our results.

Our analysis did not address risks from attribute disclosure. As noted earlier, there are no known attribute disclosure attacks on health data, and the HIPAA Privacy Rule does not require the management of attribute disclosure. This makes it difficult to determine what acceptable risk standards for attribute disclosure might be. Nevertheless, it would be appropriate to develop acceptable standards for managing attribute disclosure for future data releases.

Acknowledgments

We wish to thank Dr Jim King and Dr Carole Gentile for their advice on aspects of the work presented here. This work was funded by the Heritage Provider Network as part of the preparations for the Heritage Health Prize.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Technical description of the de-identification methods.

[PDF File (Adobe PDF File), 183KB - jmir_v14i1e33_app1.pdf]

References

1. Fienberg SE. Sharing statistical data in the biomedical and health sciences: ethical, institutional, legal, and professional dimensions. Annu Rev Public Health 1994;15:1-18. [doi: 10.1146/annurev.pu.15.050194.000245] [Medline: 8054076]
2. Sztompka P. Trust in science: Robert K Merton's inspirations. J Classical Sociol 2007;7(2):211-220. [doi: 10.1177/1468795x07078038]
3. Fienberg SE, Martin ME, Straf ML. Sharing Research Data. Washington, DC: National Academy Press; 1985.
4. Sieber JE. Data sharing: defining problems and seeking solutions. Law Hum Behav 1988;12(2):199-206. [doi: 10.1007/BF01073128]
5. Hedrick T. Justifications for the sharing of social science data. Law Hum Behav 1988;12(2):163-171. [doi: 10.1007/BF01073124]
6. Stanley B, Stanley M. Data sharing: the primary researcher's perspective. Law Hum Behav 1988;12(2):173-180. [doi: 10.1007/BF01073125]
7. Kirwan JR. Making original data from clinical studies available for alternative analysis. J Rheumatol 1997 May;24(5):822-825. [Medline: 9150066]
8. Hrynaszkiewicz I, Altman DG. Towards agreement on best practice for publishing raw clinical trial data. Trials 2009;10:17 [FREE Full text] [doi: 10.1186/1745-6215-10-17] [Medline: 19296844]
9. Piwowar HA, Day RS, Fridsma DB. Sharing detailed research data is associated with increased citation rate. PLoS One 2007;2(3):e308 [FREE Full text] [doi: 10.1371/journal.pone.0000308] [Medline: 17375194]
10. Tananbaum G. Adventures in open data. Learn Publ 2008;21(2):154-156. [doi: 10.1087/095315108x254485]
11. Murray-Rust P. Open data in science. Ser Rev 2008;34(1):52-64. [doi: 10.1016/j.serrev.2008.01.001]
12. Perry CM. Archiving of publicly funded research data: a survey of Canadian researchers. Gov Inf Q 2008;25(1):133-148. [doi: 10.1016/j.giq.2007.04.008]
13. Duncan GT, Jabine TB, de Wolf VA. Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics. Washington, DC: National Academy Press; 1993.
14. Heritage Provider Network Health Prize. 2011. Improve Healthcare, Win $3,000,000 URL: http://www.heritagehealthprize.com/ [WebCite Cache ID 65IuEDAsc]
15. Kaggle Inc. 2010. Predict HIV Progression URL: http://www.kaggle.com/c/hivprogression [WebCite Cache ID 65IuQp4ju]

16. Predictive Modeling Group, IBM Research Center. 2009. 2009 INFORMS Data Mining Contest URL: http://pages.stern.nyu.edu/~cperlich/INFORMS/ [WebCite Cache ID 65IujpMTA]
17. Phung H. Practice Fusion. 2012. Disease Control App for Doctors Wins Health 2. 0 Data Challenge; Analyze This! URL: http://www.practicefusion.com/pages/pr/disease-control-application-for-doctors-ins-health-2.0-data-challenge.html [accessed 2012-02-08] [WebCite Cache ID 65IwOTkMW]
18. Practice Fusion. Microsoft. 2010. Practice Fusion Medical Research Data URL: https://datamarket.azure.com/dataset/8c30d4d3-0846-4b08-9991-104958f24ca7 [accessed 2012-02-08] [WebCite Cache ID 65IwUjN7c]
19. El Emam K, Paton D, Dankar F, Koru G. De-identifying a public use microdata file from the Canadian national discharge abstract database. BMC Med Inform Decis Mak 2011;11:53 [FREE Full text] [doi: 10.1186/1472-6947-11-53] [Medline: 21861894]
20. Obama B. The White House. 2009. Transparency and Open Government: Memorandum for the Heads of Executive Departments and Agencies URL: http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment [accessed 2012-02-08] [WebCite Cache ID 65IwdAdyQ]
21. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. PLoS One 2011;6(12):e28071 [FREE Full text] [doi: 10.1371/journal.pone.0028071] [Medline: 22164229]
22. Association AM. Cpt 2006 Professional (Cpt / Current Procedural Terminology (Professional Edition)). Chicago, IL: American Medical Association; 2006.
23. World Health Organization. Centers for Disease Control and Prevention. 2011 Sep 23. International Classification of Diseases, 9th revision, clinical modification (ICD-9-CM) URL: http://www.cdc.gov/nchs/icd/icd9cm.htm [accessed 2012-02-13] [WebCite Cache ID 65QRE7VgL]
24. International Organization for Standardization. ISO/TS 25237:2008: Health Informatics: Pseudonymization. Geneva: International Organization for Standardization; 2009.
25. El Emam K. Heuristics for de-identifying health data. IEEE Secur Priv 2008;6(4):58-61. [doi: 10.1109/MSP.2008.84]
26. Tamersoy A, Loukides G, Denny J, Malin B. Anonymization of administrative billing codes with repeated diagnosis through censoring. In: AMIA Annu Symp Proc. 2010 Presented at: AMIA Annual Symposium; Nov 13-17, 2010; Washington, DC, USA.
27. Eguale T, Bartlett G, Tamblyn R. Rare visible disorders/diseases as individually identifiable health information. In: AMIA Annu Symp Proc. 2005 Presented at: AMIA Annual Symposium; October 22-26, 2005; Washington, DC, USA.
28. El Emam K, Dankar FK. Protecting privacy using k-anonymity. J Am Med Inform Assoc 2008;15(5):627-637 [FREE Full text] [doi: 10.1197/jamia.M2716] [Medline: 18579830]
29. Dankar F, El Emam K. A method for evaluating marketer re-identification risk. In: Proceedings of the 2010 EDBT/ICDT Workshops. 2010 Presented at: 3rd International Workshop on Privacy and Anonymity in the Information Society; Mar 22-26, 2010; Lausanne, Switzerland. [doi: 10.1145/1754239.1754271]
30. Xu Y, Fung B, Wang K, Fu A, Pei J. Publishing sensitive transactions for itemset utility. In: Proceedings of the Eighth IEEE International Conference on Data Mining. 2008 Presented at: ICDM '08; Dec 15-19, 2008; Pisa, Italy. [doi: 10.1109/ICDM.2008.98]
31. Xu Y, Wang K, Fu A, Yu P. Anonymizing transaction databases for publication. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. 2008 Presented at: KDD '08; Aug 24-27, 2008; Las Vegas, NV, USA. [doi: 10.1145/1401890.1401982]
32. Terrovitis M, Mamoulis N, Kalnis P. Privacy-preserving anonymization of set-valued data. Proceedings VLDB Endowment 2008;1(1):115-125.
33. Sweeney L. Computational Disclosure Control: A Primer on Data Privacy Protection dissertation. Boston, MA: Massachusetts Institute of Technology; 2001.
34. El Emam K, Dankar F, Vaillancourt R, Roffey T, Lysyk M. Evaluating patient re-identification risk from hospital prescription records. Can J Hosp Pharm 2009;62(4):307-319.
35. Benitez K, Malin B. Evaluating re-identification risks with respect to the HIPAA privacy rule. J Am Med Inform Assoc 2010;17(2):169-177 [FREE Full text] [doi: 10.1136/jamia.2009.000026] [Medline: 20190059]
36. Consumer-Purchaser Disclosure Project. The State Experience in Health Quality Data collection. Washington, DC: National Partnership for Women & Families; 2004 May. The State Experience in Health Quality Data collection URL: http://healthcaredisclosure.org/links/files/DataCollection.pdf [accessed 2012-02-08] [WebCite Cache ID 65J49ClNk]
37. El Emam K, Mercer J, Moreau K, Grava-Gubins I, Buckeridge D, Jonker E. Physician privacy concerns when disclosing patient data for public health purposes during a pandemic influenza outbreak. BMC Public Health 2011;11:454 [FREE Full text] [doi: 10.1186/1471-2458-11-454] [Medline: 21658256]
38. El Emam K, Dankar FK, Issa R, Jonker E, Amyot D, Cogo E, et al. A globally optimal k-anonymity method for the de-identification of health data. J Am Med Inform Assoc 2009;16(5):670-682 [FREE Full text] [doi: 10.1197/jamia.M3144] [Medline: 19567795]
39. Samarati P, Sweeney L. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalisation and suppression. Technical Report SRI-CSL-98-04. Menlo Park: SRI International; 1998. URL: http://www.csl.sri.com/papers/sritr-98-04/ [accessed 2012-02-08] [WebCite Cache ID 65J4VCiZw]

40. Samarati P. Protecting respondents' identities in microdata release. IEEE Trans Knowl Data Eng 2001;13(6):1010-1027. [doi: 10.1109/69.971193]
41. Sweeney L. k-Anonymity: a model for protecting privacy. Int J Uncertain Fuzziness Knowl Based Syst 2002;10(5):557-570. [doi: 10.1142/S0218488502001648]
42. Ciriani V, De Capitani di Vimercati S FS, Samarati P. k-Anonymity. In: Yu T, Jajodia S, editors. Secure Data Management in Decentralized Systems. New York, NY: Springer Science+Business Media, LLC; 2007.
43. Department of Health and Human Services. Federal Register. 2000 Dec 28. Standards for Privacy of Individually Identifiable Health Information URL: http://aspe.hhs.gov/admnsimp/final/PvcFR06.txt [accessed 2010-10-29] [WebCite Cache ID 5tqU5GyQX]
44. Department of Health and Human Services. Federal Register. 2000 Dec 28. Standards for Privacy of Individually Identifiable Health Information URL: http://aspe.hhs.gov/admnsimp/final/PvcFR05.txt [accessed 2010-10-29] [WebCite Cache ID 5tqULb7hT]
45. Cohn SP. Report to the Secretary of the US Department of Health and Human Services on Enhanced Protections for Uses of Health Data: A Stewardship Framework for "Secondary Uses" of Electronically Collected and Transmitted Health Data.: National Committee on Vital and Health Statistics; 2007 Dec 19. URL: http://ncvhs.hhs.gov/071221lt.pdf [accessed 2012-02-08] [WebCite Cache ID 65J5iv9O6]
46. Sweeney L. Data sharing under HIPAA: 12 years later. 2010 Presented at: Workshop on the HIPAA Privacy Rule's De-Identification Standard; Mar 8-9, 2010; Washington, DC, USA.
47. Cancer Care Ontario. Data Use & Disclosure Policy. Toronto: Cancer Care Ontario; 2005.
48. Health Quality Council. Security and Confidentiality Policies and Procedures.: Saskatchewan Health Quality Council; 2004. URL: http://www.hqc.sk.ca [accessed 2012-02-13] [WebCite Cache ID 65QRyUKXU]
49. Health Quality Council. Privacy Code. Saskatoon: Health Quality Council; 2004. URL: http://www.hqc.sk.ca/portal.jsp?G/1G3Jc1jRsvRhvh6/qyPDBIzBf0QfLQkUwK4QBZaJuIaTF3xsBNgIzOVcA+lmY4 [accessed 2012-02-02] [WebCite Cache ID 659ZIaTfD]
50. Manitoba Center for Health Policy. Privacy Code. Winnipeg: University of Manitoba; 2002 Nov. URL: http://umanitoba.ca/faculties/medicine/units/mchp/media_room/media/MCHP_privacy_code.pdf [accessed 2012-02-02] [WebCite Cache ID 659ZOOJTj]
51. Subcommittee on Disclosure Limitation Methodology, Federal Committee on Statistical Methodology. Statistical Policy Working Paper 22: Report on Statistical Disclosure Limitation Methodology. Washington, DC: Federal Office of Management and Budget; 1994 May. URL: http://www.fcsm.gov/working-papers/wp22.html [accessed 2012-02-08] [WebCite Cache ID 65J7QfbW2]
52. Statistics Canada. Therapeutic Abortion Survey. 2006. URL: http://www.statcan.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3209&lang=en&db=IMDB&dbg=f&adm=8&dis=2#b9 [accessed 2012-02-08] [WebCite Cache ID 5VkcHLeQw]
53. Information and Privacy Commissioner of British Columbia. Order No 261-1998: Inquiry re: A Media Request for Access to Records Regarding the Administration of the Medication Ritalin by School District Staff to Elementary School Students. Victoria: Office of the Information and Privacy Commissioner, Province of British Columbia; 1998 Dec 9. URL: http://www.oipc.bc.ca/orders/1998/Order261.html [accessed 2012-02-02] [WebCite Cache ID 659ZVCaVd]
54. Information and Privacy Commissioner of Ontario. Order P-644, Appeal P-9300524, Ministry of Health. Toronto: Office of the Information and Privacy Commissioner of Ontario; 1994 Mar 14. URL: http://www.ipc.on.ca/images/Findings/Attached_PDF/P-644.pdf [accessed 2009-08-05] [WebCite Cache ID 5inrVJyQp]
55. Alexander LA, Jabine TB. Access to social security microdata files for research and statistical purposes. Soc Secur Bull 1978 Aug;41(8):3-17. [Medline: 715640]
56. Ministry of Health and Long-Term Care. Corporate Policy 3-1-21. Toronto, ON: Ontario Ministry of Health and Long-Term Care; 1984.
57. de Waal A, Willenborg L. A view on statistical disclosure control for microdata. Surv Methodol 1996;22(1):95-103.
58. Office of the Privacy Commissioner of Quebec (CAI). Chenard v Ministere de l'agriculture, des pecheries et de l'alimentation (141). Quebec City: Commission d'accès à l'information; 1997.
59. Seastrom MM. NCES Statistical Standards. Source: National Center for Education Statistics, Institute of Education Sciences; 2003. URL: http://nces.ed.gov/statprog/2002/stdtoc.asp [accessed 2012-02-02] [WebCite Cache ID 659Zezo3Z]
60. Arkansas HIV/AIDS Surveillance Section. Arkansas HIV/AIDS Data Release Policy. Little Rock: Arkansas Department of Health; 2010 Mar. URL: http://www.healthy.arkansas.gov/programsServices/healthStatistics/Documents/STDSurveillance/Datadeissemination.pdf [accessed 2012-02-08] [WebCite Cache ID 65J8b3ie3]
61. Centers for Medicare & Medicaid Services. BSA Inpatient Claims PUF. 2011. BSA Inpatient Claims PUF URL: http://www.cms.gov/BSAPUFS/03_Inpatient_Claims.asp [accessed 2012-02-08] [WebCite Cache ID 65J8juTAp]
62. Escobar GJ, Greene JD, Scheirer P, Gardner MN, Draper D, Kipnis P. Risk-adjusting hospital inpatient mortality using automated inpatient, outpatient, and laboratory databases. Med Care 2008 Mar;46(3):232-239. [doi: 10.1097/MLR.0b013e3181589bb6] [Medline: 18388836]

63. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [Medline: 3558716]
64. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005 Nov;43(11):1130-1139. [Medline: 16224307]
65. Liu J, Wang K. Anonymizing transaction data by integrating suppression and generalization. In: Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science. Berlin: Springer-Verlag; 2010 Presented at: PAKDD 2010; Jun 21-24, 2010; Hyderabad, India. [doi: 10.1007/978-3-642-13657-3_20]
66. He Y, Naughton J. Anonymization of set-valued data via top-down, local generalization. Proceedings VLDB Endowment 2009;2(1):934-945.
67. Davison A, Hinkley D. Bootstrap methods and their application. Cambridge: Cambridge University Press; 1997.

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
CPT: Current Procedural Terminology
HHP: Heritage Health Prize
HIPAA: Health Insurance Portability and Accountability Act
HPN: Heritage Provider Network
ICD-9-CM: International Classification of Diseases, 9th revision, clinical modification
LOLA: longitudinal optimal lattice anonymization
LOS: length of stay
OLA: optimal lattice anonymization
PCG: primary condition group
SID: State Inpatient Database

Edited by G Eysenbach; submitted 20.11.11; peer-reviewed by T Gal; comments to author 12.12.11; revised version received 14.01.12; accepted 17.01.12; published 27.02.12.

Please cite as:
El Emam K, Arbuckle L, Koru G, Eze B, Gaudette L, Neri E, Rose S, Howard J, Gluck J
De-identification Methods for Open Health Data: The Case of the Heritage Health Prize Claims Dataset
J Med Internet Res 2012;14(1):e33
URL: http://www.jmir.org/2012/1/e33/
doi:10.2196/jmir.2001
PMID:22370452

Original Paper

Real-time Prescription Surveillance and its Application to Monitoring Seasonal Influenza Activity in Japan

Tamie Sugawara[1], PhD; Yasushi Ohkusa[1], PhD; Yoko Ibuka[2], PhD; Hirokazu Kawanohara[3], BS; Kiyosu Taniguchi[1], PhD, MD; Nobuhiko Okabe[1], PhD, MD

[1]National Institute of Infectious Diseases, Infectious Disease Surveillance Center, Tokyo, Japan

[2]Hitotsubashi University, Tokyo, Japan

[3]EM Systems Co., Ltd, Tokyo, Japan

Corresponding Author:
Tamie Sugawara, PhD
National Institute of Infectious Diseases
Infectious Disease Surveillance Center
1-23-1Toyama
Shinjuku
Tokyo, 162-8640
Japan
Phone: 81 3 5285 1111
Fax: 81 3 5285 1129
Email: tammy@nih.go.jp

Abstract

Background: Real-time surveillance is fundamental for effective control of disease outbreaks, but the official sentinel surveillance in Japan collects information related to disease activity only weekly and updates it with a 1-week time lag.

Objective: To report on a prescription surveillance system using electronic records related to prescription drugs that was started in 2008 in Japan, and to evaluate the surveillance system for monitoring influenza activity during the 2009–2010 and 2010–2011 influenza seasons.

Methods: We developed an automatic surveillance system using electronic records of prescription drug purchases collected from 5275 pharmacies through the application service provider's medical claims service. We then applied the system to monitoring influenza activity during the 2009–2010 and 2010–2011 influenza seasons. The surveillance system collected information related to drugs and patients directly and automatically from the electronic prescription record system, and estimated the number of influenza cases based on the number of prescriptions of anti-influenza virus medication. Then it shared the information related to influenza activity through the Internet with the public on a daily basis.

Results: During the 2009–2010 influenza season, the number of influenza patients estimated by the prescription surveillance system between the 28th week of 2009 and the 12th week of 2010 was 9,234,289. In the 2010–2011 influenza season, the number of influenza patients between the 36th week of 2010 and the 12th week of 2011 was 7,153,437. The estimated number of influenza cases was highly correlated with that predicted by the official sentinel surveillance ($r = .992$, $P < .001$ for 2009–2010; $r = .972$, $P < .001$ for 2010–2011), indicating that the prescription surveillance system produced a good approximation of activity patterns.

Conclusions: Our prescription surveillance system presents great potential for monitoring influenza activity and for providing early detection of infectious disease outbreaks.

(J Med Internet Res 2012;14(1):e14) doi:10.2196/jmir.1881

KEYWORDS

Surveillance; influenza; real-time surveillance; prescriptions; pharmacy; anti-influenza virus; automatic surveillance; early response

Introduction

In Japan, the official sentinel surveillance reports the number of influenza patients per health care provider after collecting information from approximately 5000 clinics and hospitals. The intensity of influenza activity is assessed according to the number of influenza patients per clinic or hospital. Influenza is regarded as highly active if the ratio exceeds 1. In 2009, the number of patients per clinic or hospital approached 1 in the 32nd week, earlier than in any of the preceding 10 years, mainly because of the influenza pandemic A (H1N1), which started in April 2009 [1]. Accordingly, the vast majority of the reported cases were H1N1 novel influenza [1]. The number of influenza patients per health care provider declined below 1 in the 13th week of 2010. The total number of weeks during which influenza was highly active was 29, a longer active period than in any of the prior 10 years. In 2010, the reported number of influenza patients per clinic or hospital exceeded 1 in the 50th week [2]; a second peak week was detected in March 2011. Because of these irregular patterns of influenza activity, it is necessary that both policy makers and clinicians follow influenza activity closely to implement effective control of an influenza outbreak throughout the year.

Syndromic surveillance is a useful tool for seasonal influenza monitoring [3]. In Japan, the official sentinel surveillance of infectious diseases is implemented by the National Institute of Infectious Diseases. It reports the estimated number of influenza patients weekly as the *Infectious Diseases Weekly Report* [2]. The official sentinel surveillance collects the number of influenza cases from approximately 5000 hospitals and clinics all over the country and then estimates the number of influenza patients based on the reported cases [4]. The entire process of collecting information from health care providers, estimating the number of clinical influenza cases, and reporting them to the public usually takes 7–10 days. Furthermore, the cases are reported by health care providers as a weekly aggregate number. Some diseases spread rapidly, and the weekly aggregates might not provide sufficiently detailed information reflecting the complete character of disease activity. In addition, the official sentinel surveillance updates influenza activity less frequently during major holidays. In Japan, seasonal influenza activity usually starts to become active during the New Year holidays. Constant monitoring and reporting of activity during that period is necessary.

Syndromic surveillance is in widespread use for monitoring diseases, but usage of prescription drug sales as a source of information is fairly limited. In the United States, the most common source of syndromic surveillance reported by health officials is emergency department visits (84%), followed by outpatient clinic visits (49%) and over-the-counter medication sales (44%); less than 10% of health departments reported prescription medications as a source [3]. In the context of influenza, emergency department surveillance is used to monitor the impact of influenza by age [5]. For more rapid feedback, the Web recently has become a powerful tool for syndromic surveillance [6]. For example, health surveillance using a Web-based self-reporting daily questionnaire is applied to monitor influenza activities [7]. Google Flu Trends, a Web-based surveillance, tracks the rate of influenza using query logs [8]. In addition to monitoring disease activities, syndromic surveillance helps monitor bioterrorism-related disease [9] or health consequences of natural events [10].

Real-time information related to influenza activity is fundamentally important for better preparation of countermeasures against a sudden increase of influenza activity. Therefore, daily updates of influenza activity are indispensable for improved understanding and control of an influenza epidemic. We developed an automatic real-time prescription surveillance system with the collaboration of EM Systems Co. Ltd. (Tokyo, Japan) to provide timely information related to a disease outbreak. We applied the surveillance system to monitor influenza activity during the 2009–2010 and 2010–2011 influenza seasons to examine the magnitude and trajectory of an outbreak more closely and to share that information with public health authorities, as well as participating pharmacies.

We used prescription drug purchase data for surveillance of influenza activity for three reasons. First, prescribing anti-influenza drugs such as oseltamivir or zanamivir is a common clinical practice for diagnosed influenza cases in Japan. Japan has the highest annual level of oseltamivir usage in the world [11]. Therefore, prescription drugs can serve as a good indicator of the overall number of influenza patients. Physicians often perform rapid influenza diagnostic tests on patients who have a fever or report other influenza-like symptoms. If the test result is positive or, alternatively, if the physician clinically diagnoses influenza even when the test result is negative, then anti-influenza drugs are often prescribed. This contrasts to practices in some other developed countries, where anti-influenza drugs are recommended for those who are at high risk [12-14] or who have severe conditions from influenza infections [13,14]. In such circumstances, surveillance of prescriptions of anti-influenza drugs would trace influenza patients with severe symptoms [15].

Second, many pharmacies have adopted the electronic prescription record system (EPRS), which enables automatic, continuous, and constant information collection, and real-time analysis of prescriptions and patients. In Japan, the utilization rate of the EPRS among pharmacies was 99.0% in 2009 [16]. Japan also has a high rate of outpatient or office-based clinician visits in cases where people feel ill [17], partly because of the universal health insurance system. Therefore, one might infer that the number of influenza patients collected through the EPRS would closely approximate the number of symptomatic influenza patients.

Third, in contrast to the United States or Taiwan [18], in Japan electronic medical record (EMR) systems are not yet well established. In the United States, surveillance for influenza activity is based on data on outpatient visits along with data related to sales of over-the-counter drugs, school absenteeism, and ambulatory care encounters [3,9,19-21]. Surveillance for influenza activity using the EMR has been intensively discussed and widely applied [22-24]. By contrast, the Survey of Medical Institutions by the Ministry of Health, Labour and Welfare in Japan showed that the share of health care providers using EMRs

was just over 10% in 2008, or 948 hospitals (10.8% of all hospitals) and 12,939 clinics (13.1% of all clinics) [25].

We developed the surveillance system to collect the number of prescriptions together with patients' characteristics from the EPRS automatically, to analyze the data simultaneously to estimate the number of influenza cases, and then to provide real-time information of influenza activity to health care providers and policy makers. The system was tested for a limited time at the G8 Summit meeting in Toyako, Hokkaido in July 2008 for 1 month [26]. The present report summarizes details of our prescription surveillance system and presents an evaluation of its performance in the first two influenza seasons, those of 2009–2010 and 2010–2011, since the start of the nationwide operation of the system. The evaluation of surveillance performance, particularly outbreak detection performance, is challenging and few studies conduct such analyses [27]. A study showed that weekly variation in visits for lower respiratory tract infections approximated the national mortality data for pneumonia and influenza [28]. Similarly, our retrospective evaluation analyzed how closely the estimates of influenza cases followed the trajectory of influenza epidemics reported by two other sources.

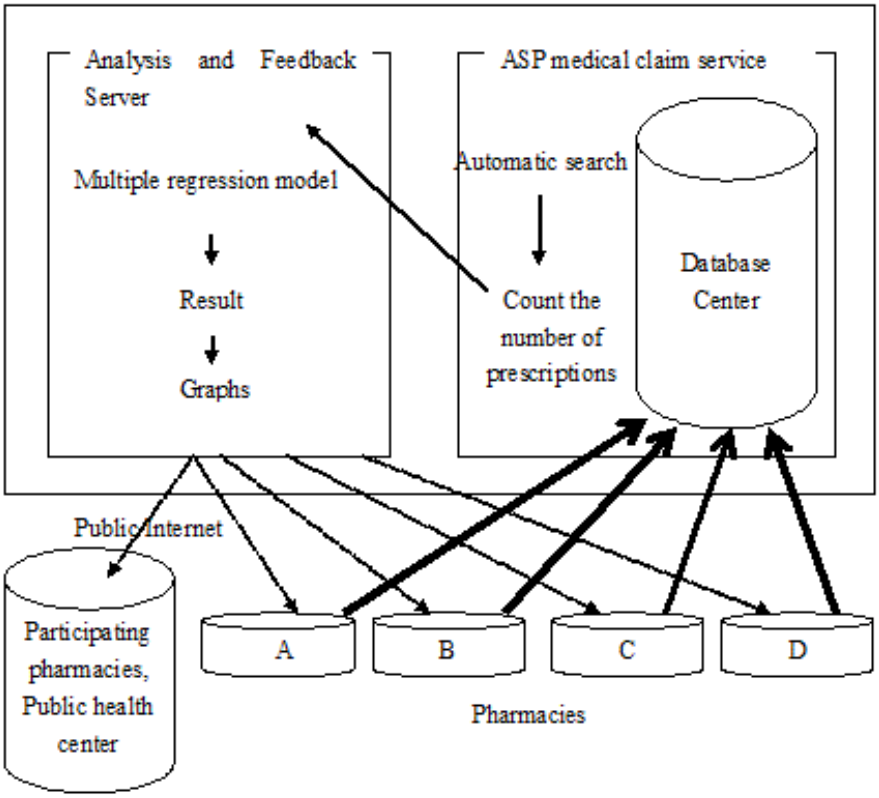
Methods

Prescription Surveillance

We started collecting and analyzing data related to prescriptions automatically through the application service provider of the EPRS in April 2009 (Figure 1 [29]). As of March 2011, the number of participating pharmacies was 5275. In the application service provider, data related to prescriptions from all participating pharmacies were collected and deposited in a single server, making the data collection secure, efficient, and nearly cost-free. Medications covered by the surveillance system included drugs for relief of fever and pain, drugs for common colds, antibiotics, and antiviral drugs including anti-influenza virus drugs and antivaricella-zoster virus drugs. The current study specifically addressed prescriptions for anti-influenza virus medication. The neuraminidase inhibitors oseltamivir, zanamivir, and laninamivir were included, but amantadine was excluded because it is not commonly prescribed for influenza in Japan.

The original prescriptions contain information related to patients' sociodemographic and social security information, as well as the health care providers' information. The automatic surveillance system aggregated the number of prescriptions for each type of drug and provided tabulations by age and by geography at both national and prefectural levels. The number of influenza patients was then estimated from the aggregated number of prescriptions for anti-influenza drugs by adjusting the number of prescriptions for anti-influenza drugs with the proportion of participating pharmacies and of prescriptions purchased through pharmacies. The analysis and estimation were conducted overnight and the report of the analysis was sent automatically at 7:00 AM on the next day to the registered recipients, including participating pharmacies and public health authorities. In addition, figures showing the number of prescriptions for each type of drug and of the estimated number of patients were created and posted on the website for public access.

Figure 1. Prescription surveillance. Pharmacies A–D use the application service provider's (ASP) medical claims service. All data are stored in a central database. The surveillance system automatically counts oseltamivir, zanamivir, and laninamivir prescriptions at the data center. The information is analyzed using multiple regression models. The results are presented as figures and tables and feedback to participating pharmacies as well as public health authorities.



Performance Evaluation

We evaluated our surveillance system from two perspectives for the 2009–2010 and 2010–2011 influenza seasons. First, we compared the estimated number of influenza patients with the estimates provided by the official sentinel surveillance [2]. The official sentinel surveillance estimates the number of influenza patients based on the number of influenza patients reported by 5000 health care providers, including 3000 pediatricians, in Japan. We chose the evaluation period to include the period when influenza activity was high for the 2009–2010 influenza season. The epidemiological threshold of seasonal influenza activity is determined by the number of influenza patients per hospital or clinic. If the ratio is equal to or greater than 1 based on the official sentinel surveillance, activity is *high* by the definition that is accepted and widely used throughout Japan [2]. This corresponds to the period between the 28th week of 2009 (the week starting on July 6, 2009) and the 12th week of 2010 (the week starting on March 21, 2010) for the 2009–2010 influenza season. For the 2010–2011 season, the performance was evaluated between the 36th week of 2010 (the week starting on September 6, 2010) and the 12th week of 2011 (the week starting on March 21, 2011). Second, for the 2009–2010 influenza season, we also compared our estimates with the number of influenza patients estimated by the Gifu Medical Association, where the total number of influenza patients in the prefecture was calculated and reported publicly [29]. The number of influenza patients in Gifu Prefecture was surveyed during November 16–22, 2009 by the local public health authority as a response to the A/H1N1 influenza pandemic. A survey questionnaire asking for the number of influenza patients who visited health care providers was sent to all hospitals and clinics located within the prefecture (total of 1677 health providers); 1033 providers responded to the survey (response rate 61.6%) [29].

The Internal Review Board at the National Institute of Infectious Diseases approved the current study (approval number 57, "Development and application of real-time surveillance system to monitor syndromic and symptomatic cases using electronic record system").

Results

For the 2009–2010 influenza season, the total number of influenza patients estimated by the prescription surveillance system between the 28th week of 2009 and the 12th week of 2010 was 9,234,289 (Table 1). The largest number of influenza patients, 234,519, was reported on November 24, 2009. For the 2010–2011 influenza season, the number of influenza patients between the 36th week of 2010 and the 12th week of 2011 was 7,153,437 (Table 1). The largest number of influenza patients, 230,288, was reported on January 24, 2011. The official sentinel

surveillance estimated the total number of patients for the same periods as 20,660,000 (95% confidence interval 20,460,000–20,860,000) for the 2009–2010 and 13,680,000 (95% confidence interval 13,350,000–14,010,000) for the 2010–2011 influenza seasons [2], indicating that the sentinel estimates were approximately double our estimates.

Table 1. Number of influenza cases estimated by the prescription surveillance, the official sentinel surveillance, and the Gifu Medical Association in Gifu Prefecture, 2009–2010 and 2010–2011 influenza seasons[a]

	2009–2010 influenza season: July 6, 2009–March 28, 2010 (28th week 2009–12th week 2010)	2010–2011 influenza season: September 6, 2010–March 27, 2011 (36th week 2010–12th week 2011)
Estimate by the prescription surveillance	9,234,289	7,153,437
Estimate by the official sentinel surveillance	20,660,000	13,680,000
Adjusted estimation by the survey in Gifu Prefecture	9,931,200	Not applicable[b]

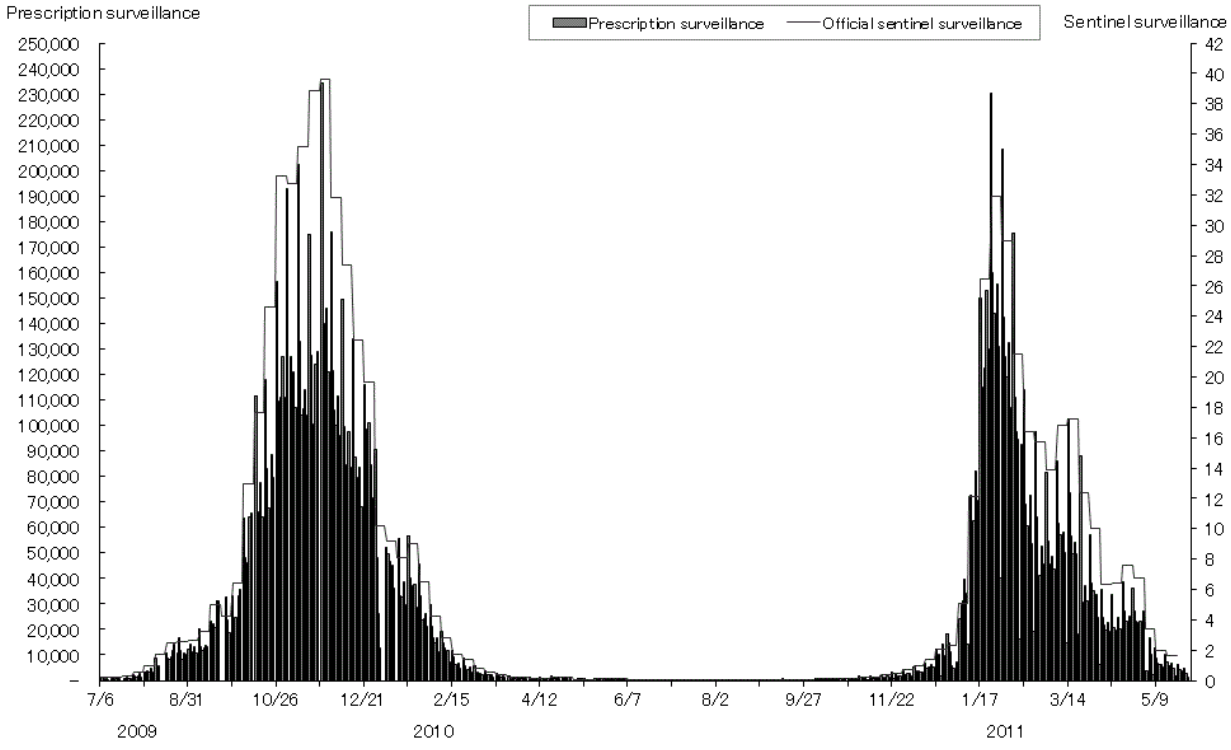
[a] Sources: the official sentinel surveillance [2]; Kawai et al [29].

[b] Adjusted estimation by the survey in Gifu Prefecture is shown only for the 2009–2010 influenza season because the data are available only for that year.

Pearson correlation coefficient (r) of time-series data on influenza patients between our estimates and the official sentinel estimate was .992 ($P < .001$) for the 2009–2010 influenza season, and .972 ($P < .001$) for the 2010–2011 influenza season (see Figure 2). A similar analysis was conducted at the prefecture level. The correlation was.950 or greater in 33 prefectures, .900–.949 in 5 prefectures, and .770–.899 in 8 prefectures. The correlation was the lowest in Akita Prefecture ($r = .689$).

The estimated number of influenza cases in the 2009–2010 influenza season was also compared with that ascertained from the survey of the number of influenza patients at all clinics and hospitals conducted in Gifu Prefecture. The estimated number from the survey collection in the prefecture based on the prescription surveillance was 127,568, whereas the number of influenza cases reported by the survey conducted by Gifu Medical Association was 132,474. The official sentinel surveillance estimated the number as 277,890.

Figure 2. Number of influenza cases, 2009–2011, estimated by the prescription surveillance and reported by the official sentinel surveillance. The estimated number of influenza cases by prescription surveillance was calculated based on the number of oseltamivir, zanamivir, and laninamivir prescriptions adjusted by the proportion of participating pharmacies and extramural dispensing percentage. See text for details. The reported number by the official sentinel surveillance shows the number of influenza patients per clinic or hospital, calculated with the reported number of influenza patients from 5000 sentinel clinics and hospitals.



Discussion

Our analyses showed that the time-series pattern of influenza activity reported by the prescription surveillance system in the first two influenza seasons was highly correlated with the pattern reported by the official sentinel surveillance, showing that pharmacy surveillance can be a good indicator of influenza activity in Japan. Although the estimated number of influenza patients was double that of the official sentinel surveillance, it was close to the estimate by Gifu Prefecture, where the total number of influenza patients was collected in a survey.

The significance of our prescription surveillance is threefold. First, the syndromic surveillance system collected, analyzed, and reported data related to influenza patients simultaneously. Therefore, clinicians and policy makers were able to obtain the estimated number of influenza patients of the previous day. This meant that the estimates were available 1 week ahead of those reported by the official sentinel surveillance, enabling predictions of influenza activity for the immediately following week. This was particularly important at the outset of a seasonal epidemic, when the trajectory of a quickly spreading disease would have changed. Though the Google Flu Trends tool, another real-time surveillance, has been shown to perform well in the United States [8] and European countries [30], the results may be sensitive to variations in patients' behavior across countries.

Second, our prescription surveillance was national and observed regional variations in influenza activity at the prefecture level, although the precision of surveillance varied somewhat between prefectures. This provided helpful information to public health services to plan for the allocation of medical, pharmaceutical, and human resources for influenza control, shifting limited resources to the most affected regions.

Third, our surveillance runs constantly, maintaining the method of counting and estimating influenza cases at all times, and thus we were able to obtain the complete trajectory of the influenza pandemic in the 2009–2010 season. Initially during the pandemic, the law required hospitals and clinics to report all influenza cases, but that practice was terminated on July 24, 2009, after which activity was tracked only by the official sentinel surveillance.

Our surveillance system also promises great potential for future application to the early detection of an infectious disease outbreak or bioterrorism attack, which could happen potentially anywhere at any time. When we started operating a prescription surveillance system in 2009, all other surveillance systems running in Japan covered only specific regions of the country for practical reasons [31]. Furthermore, because influenza outbreaks do not necessarily occur during winter, the time that is covered by the sentinel surveillance, continuous monitoring of influenza activity is necessary to detect outbreaks early in their course. Our automatic prescription surveillance system uses the same standard for detection of a disease outbreak and runs continuously, providing an important complementary role in support of existing surveillance systems in Japan.

If EMRs were widely kept, then information related to influenza patients could be collected even faster and possibly more accurately. However, the share of health care providers that have adopted the EMR system was slightly above 10%. Under such circumstances, purchases of anti-influenza drugs can serve as an alternative indicator of influenza activity.

Limitations to this study exist. First, the total number of influenza cases was estimated as almost half of the estimate based on the official sentinel surveillance, although it approximated estimates based on a survey collecting the total number of influenza cases in Gifu Prefecture. One reason for this gap might lie in the choice of health care providers participating in the official sentinel surveillance. The sentinel health care providers have, on average, a larger number of patients than others, potentially resulting in an overestimation of the overall number of influenza patients. Second, anti-influenza drugs are also prescribed for prophylaxis in addition to treatment, which might engender overestimation of the total number of influenza cases. However, in Japan the preventive usage of oseltamivir is limited to household members of influenza patients who are 65 years or older or who are high-risk individuals [32]. Third, the prophylactic usage of anti-influenza drugs for health care providers and for the public was most intensive at the beginning of the H1N1 pandemic outbreak. We did not include those prescriptions in our surveillance data because they were not prescribed through health care providers. Fourth, 60% of the prescriptions were purchased through pharmacies in 2008. The other prescriptions were purchased directly through health care providers and were not included in our surveillance [33]. This is still much higher than the rate of adoption of the EMR system in hospitals and clinics. Fifth, the participation rate of pharmacies is low, particularly in certain areas. If the number of participating pharmacies were increased, then estimating influenza cases would be possible even for smaller geographical units.

Despite these limitations, pharmacy surveillance provided an approximation of the trend of influenza activity in the first two influenza seasons after the start of its nationwide operation. It provided both clinicians and policy makers with helpful real-time information related to influenza activity. Our pharmacy surveillance system has great potential for detection as well as for monitoring of infectious disease outbreaks in the population and in cases of significant political or cultural events.

Acknowledgments

This research was financially supported by a Health and Labour Sciences Research Grant from the Ministry of Health, Labour and Wealth, "Research for practical application of early detection system for health risk," headed by the second author, Dr Yasushi Ohkusa. Dr Yoko Ibuka received a grant from the Abe Fellowship Program administered by the Social Science Research Council

and the American Council of Learned Societies in cooperation with and with funds provided by the Japan Foundation Center for Global Partnership.

Conflicts of Interest

None declared.

References

1. Shimada T, Gu Y, Kamiya H, Komiya N, Odaira F, Sunagawa T, et al. Epidemiology of influenza A(H1N1)v virus infection in Japan, May-June 2009. Euro Surveill 2009 Jun 18;14(24) [FREE Full text] [Medline: 19555600]
2. National Institute of Infectious Diseases. 2011. Infectious Diseases Weekly Report: Trend Graph [in Japanese] URL: http://idsc.nih.go.jp/idwr/index.html [accessed 2011-06-30] [WebCite Cache ID 5zou5OUNU]
3. Buehler JW, Sonricker A, Paladini M, Sope P, Mostashar F. Syndromic surveillance practice in the United States: findings from a survey of state, territorial, and selected local health departments. Adv Dis Surveill 2008;6(3).
4. Taniguchi K, Hashimoto S, Kawado M, Murakami Y, Izumida M, Ohta A, et al. Overview of infectious disease surveillance system in Japan, 1999-2005. J Epidemiol 2007 Dec;17 Suppl:S3-13 [FREE Full text] [Medline: 18239339]
5. Olson DR, Heffernan RT, Paladini M, Konty K, Weiss D, Mostashari F. Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City. PLoS Med 2007 Aug;4(8):e247 [FREE Full text] [doi: 10.1371/journal.pmed.0040247] [Medline: 17683196]
6. 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] [Medline: 18613747]
7. Sugiura H, Ohkusa Y, Akahane M, Sano T, Okabe N, Imamura T. Development of a web-based survey for monitoring daily health and its application in an epidemiological survey. J Med Internet Res 2011;13(3):e66 [FREE Full text] [doi: 10.2196/jmir.1872] [Medline: 21946004]
8. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
9. Buehler JW, Berkelman RL, Hartley DM, Peters CJ. Syndromic surveillance and bioterrorism-related epidemics. Emerg Infect Dis 2003 Oct;9(10):1197-1204 [FREE Full text] [Medline: 14609452]
10. Elliot AJ, Singh N, Loveridge P, Harcourt S, Smith S, Pnaiser R, et al. Syndromic surveillance to assess the potential public health impact of the Icelandic volcanic ash plume across the United Kingdom, April 2010. Euro Surveill 2010;15(23) [FREE Full text] [Medline: 20546694]
11. Ujike M, Shimabukuro K, Mochizuki K, Obuchi M, Kageyama T, Shirakura M, Working Group for Influenza Virus Surveillance in Japan. Oseltamivir-resistant influenza viruses A (H1N1) during 2007-2009 influenza seasons, Japan. Emerg Infect Dis 2010 Jun;16(6):926-935 [FREE Full text] [doi: 10.3201/eid1606.091623] [Medline: 20507742]
12. National Institute for Health and Clinical Excellence. 2009 Feb. Amantadine, Oseltamivir and Zanamivir for the Treatment of influenza: Review of NICE Technology Appraisal Guidance 58 (NICE Technology Appraisal Guidance 168) URL: http://www.nice.org.uk/nicemedia/live/11774/43268/43268.pdf [accessed 2011-06-30] [WebCite Cache ID 5zospD9eK]
13. Centers for Disease Control and Prevention. 2011 Jan 31. Interim Guidance on the Use of Influenza Antiviral Agents During the 2010-2011 Influenza Season URL: http://www.cdc.gov/flu/professionals/antivirals/guidance/ [accessed 2011-06-30] [WebCite Cache ID 5zot1AMbz]
14. Harper SA, Bradley JS, Englund JA, File TM, Gravenstein S, Hayden FG, Expert Panel of the Infectious Diseases Society of America. Seasonal influenza in adults and children--diagnosis, treatment, chemoprophylaxis, and institutional outbreak management: clinical practice guidelines of the Infectious Diseases Society of America. Clin Infect Dis 2009 Apr 15;48(8):1003-1032 [FREE Full text] [doi: 10.1086/598513] [Medline: 19281331]
15. Sugawara T, Ohkusa Y, Kawanohara H, Taniguchi K, Okabe N. [The real-time pharmacy surveillance and its estimation of patients in 2009 influenza A (H1N1)]. Kansenshogaku Zasshi 2011 Jan;85(1):8-15. [Medline: 21404600]
16. Ministry of Health, Labour and Welfare, Japan. 2010. Constructing an Appropriate, Stable and Efficient Healthcare Insurance System [in Japanese] URL: http://www.mhlw.go.jp/wp/seisaku/jigyou/10jisseki/dl/youshi/1-10-1.pdf [accessed 2011-06-30] [WebCite Cache ID 5zotEB0Qe]
17. Sugawara T, Ohkusa Y, Kondo M, Honda Y, Okubo I. [Research on choices of people with mild symptoms of common cold between consulting physicians and taking OTC (over-the-counter) medicine using a hypothetical question method]. Nihon Koshu Eisei Zasshi 2005 Jul;52(7):618-626. [Medline: 16130888]
18. Wu TS, Shih FY, Yen MY, Wu JS, Lu SW, Chang KC, et al. Establishing a nationwide emergency department-based syndromic surveillance system for better public health responses in Taiwan. BMC Public Health 2008;8:18 [FREE Full text] [doi: 10.1186/1471-2458-8-18] [Medline: 18201388]
19. Lazarus R, Kleinman K, Dashevsky I, Adams C, Kludt P, DeMaria A, et al. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. Emerg Infect Dis 2002 Aug;8(8):753-760 [FREE Full text] [Medline: 12141958]

20. Henning KJ. What is syndromic surveillance? MMWR Morb Mortal Wkly Rep 2004 Sep 24;53 Suppl:5-11 [FREE Full text] [Medline: 15714620]
21. Ohkusa Y, Sugiura H, Sugawara T, Taniguchi K, Okabe N. [Symptoms of outpatients as data for syndromic surveillance]. Kansenshogaku Zasshi 2006 Jul;80(4):366-376. [Medline: 16922479]
22. South BR, South BR, Chapman WW, Chapman W, Delisle S, Shen S, et al. Optimizing A syndromic surveillance text classifier for influenza-like illness: Does document source matter? AMIA Annu Symp Proc 2008:692-696. [Medline: 18999051]
23. Gundlapalli AV, Olson J, Smith SP, Baza M, Hausam RR, Eutropius LJ, et al. Hospital electronic medical record-based public health surveillance system deployed during the 2002 Winter Olympic Games. Am J Infect Control 2007 Apr;35(3):163-171. [doi: 10.1016/j.ajic.2006.08.003] [Medline: 17433939]
24. Lewis MD, Pavlin JA, Mansfield JL, O'Brien S, Boomsma LG, Elbert Y, et al. Disease outbreak detection system using syndromic data in the greater Washington DC area. Am J Prev Med 2002 Oct;23(3):180-186. [Medline: 12350450]
25. Ministry of Health, Labour and Welfare, Japan. 2011 Jun 27. Survey of Medical Institutions URL: http://www.mhlw.go.jp/english/database/db-hss/mi.html [accessed 2011-06-30] [WebCite Cache ID 5zotSWn20]
26. Ohkusa Y, Yamaguchi R, Sugiura H, Sugawara T, Yoshida M, Shimada C, et al. [2008 G8 Hokkaido Toyako Summit Meeting Syndrome Surveillance]. Kansenshogaku Zasshi 2009 May;83(3):236-244. [Medline: 19522307]
27. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. J Am Med Inform Assoc 2004 Apr;11(2):141-150 [FREE Full text] [doi: 10.1197/jamia.M1356] [Medline: 14633933]
28. Lazarus R, Kleinman K, Dashevsky I, DeMaria A, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of lower respiratory infection. BMC Public Health 2001;1:9 [FREE Full text] [Medline: 11722798]
29. Kawai N, Kawade Y, Kobayashi H, Okada S, Higuchi Y, Kawaji H, et al. Analysis of influenza activity during 2009 influenza pandemic A (H1N1) using real-time infectious disease surveillance in Gifu prefecture [in Japanese]. Syukan Nihon Iji Shinpou 2010;4487:58-64.
30. Valdivia A, Lopez-Alcalde J, Vicente M, Pichiule M, Ruiz M, Ordobas M. Monitoring influenza activity in Europe with Google Flu Trends: comparison with the findings of sentinel physician networks - results for 2009-10. Euro Surveill 2010;15(29) [FREE Full text] [Medline: 20667303]
31. Ohkusa Y, Shigematsu M, Taniguchi K, Okabe N. Experimental surveillance using data on sales of over-the-counter medications--Japan, November 2003-April 2004. MMWR Morb Mortal Wkly Rep 2005 Aug 26;54 Suppl:47-52 [FREE Full text] [Medline: 16177693]
32. Chugai Pharmaceutical Co. Ltd. 2011 Jun 30. Tamiflu 75 [in Japanese] URL: http://www.info.pmda.go.jp/go/pack/6250021M1027_1_24/ [accessed 2011-06-30] [WebCite Cache ID 5zotcTEZR]
33. Ministry of Health, Labour and Welfare, Japan. 2008. Survey of Medical Care Activities in Public Health Insurance [in Japanese] URL: http://www.mhlw.go.jp/toukei/saikin/hw/sinryo/tyosa08/index.html [accessed 2011-06-30] [WebCite Cache ID 5zotlEjcg]

Abbreviations

EMR: electronic medical record
EPRS: electronic prescription record system

Edited by G Eysenbach; submitted 30.06.11; peer-reviewed by D Zeng; comments to author 26.09.11; revised version received 25.10.11; accepted 01.11.11; published 16.01.12.

Please cite as:
Sugawara T, Ohkusa Y, Ibuka Y, Kawanohara H, Taniguchi K, Okabe N
Real-time Prescription Surveillance and its Application to Monitoring Seasonal Influenza Activity in Japan
J Med Internet Res 2012;14(1):e14
URL: http://www.jmir.org/2012/1/e14/
doi:10.2196/jmir.1881
PMID:22249906

Original Paper

eHealth Literacy: Extending the Digital Divide to the Realm of Health Information

Efrat Neter[1], PhD; Esther Brainin[1], PhD

Behavioral Sciences Department, Ruppin Academic Center, Emeq Hefer, Israel

Corresponding Author:
Efrat Neter, PhD
Behavioral Sciences Department
Ruppin Academic Center
Bait 3, Ruppin Academic Center
Emeq Hefer, 40250
Israel
Phone: 972 98981352
Fax: 972 98987604
Email: neter@ruppin.ac.il

Abstract

Background: eHealth literacy is defined as the ability of people to use emerging information and communications technologies to improve or enable health and health care.

Objective: The goal of this study was to explore whether literacy disparities are diminished or enhanced in the search for health information on the Internet. The study focused on (1) traditional digital divide variables, such as sociodemographic characteristics, digital access, and digital literacy, (2) information search processes, and (3) the outcomes of Internet use for health information purposes.

Methods: We used a countrywide representative random-digital-dial telephone household survey of the Israeli adult population (18 years and older, N = 4286). We measured eHealth literacy; Internet access; digital literacy; sociodemographic factors; perceived health; presence of chronic diseases; as well as health information sources, content, search strategies, and evaluation criteria used by consumers.

Results: Respondents who were highly eHealth literate tended to be younger and more educated than their less eHealth-literate counterparts. They were also more active consumers of all types of information on the Internet, used more search strategies, and scrutinized information more carefully than did the less eHealth-literate respondents. Finally, respondents who were highly eHealth literate gained more positive outcomes from the information search in terms of cognitive, instrumental (self-management of health care needs, health behaviors, and better use of health insurance), and interpersonal (interacting with their physician) gains.

Conclusions: The present study documented differences between respondents high and low in eHealth literacy in terms of background attributes, information consumption, and outcomes of the information search. The association of eHealth literacy with background attributes indicates that the Internet reinforces existing social differences. The more comprehensive and sophisticated use of the Internet and the subsequent increased gains among the high eHealth literate create new inequalities in the domain of digital health information. There is a need to educate at-risk and needy groups (eg, chronically ill) and to design technology in a mode befitting more consumers.

(J Med Internet Res 2012;14(1):e19) doi:10.2196/jmir.1619

KEYWORDS

eHealth literacy; digital literacy; health literacy; digital divide; health information search

Introduction

eHealth, a relatively new concept [1], refers to "the use of emerging information and communications technology to improve or enable health and health care" [2]. eHealth literacy, which includes the component of *health literacy* [3-5], effectively links health consumers to the outcomes typical of Internet use—that is, opportunities, possible harm [6], and inequalities (eg, being part of a minority or disenfranchised

group [7-9], education [10-13], age [13-15], and gender [15-17]). In the 1990s, the concern over inequalities related to the digital divide focused mainly on infrastructural access: ownership, availability, and affordability of the infrastructure [18]. The discourse on the digital divide has expanded to other concerns, shifting the emphases to patterns of access [19], usage [20,21], and online skills rather than mere access to technology [13,20,21]. eHealth literacy may constitute a second divide in the health domain [21,22].

Norman and Skinner [23] propose that eHealth literacy is "the ability to seek, find, understand and appraise health information from electronic sources and apply knowledge gained to addressing or solving a health problem." They propose that eHealth literacy encompasses 6 kinds of literacies: *traditional* (literacy and numeracy), *information*, *media*, *health*, *computer*, and *scientific*. Of these, media and computer literacies are unique attributes of the Internet context, with media being the awareness of media bias or perspective, the ability to discern both explicit and implicit meaning from media messages, and deriving meaning from media messages. The literature includes other ways in which perceived capability or efficacy was measured, but these were not specific to health information on the Internet [24-26].

Norman and Skinner [23,27] advocate matching eHealth technologies to the skills of their intended users. Such a fit can be realized by improving users' working knowledge of computers (or of the particular language or skill) to a level conducive to achieving health-related goals, as well as by designing systems with the users in mind. To further address this divide, Norman and Skinner developed an eHealth literacy scale (eHEALS) to measure eHealth literacy [27], using a sample of Canadian adolescents. They emphasized that eHealth literacy should be viewed as a process that evolves over time, while the pace of development depends on technology and context (personal, social, and environmental), rather than on a static attribute. Viewed as malleable, eHealth literacy may indeed "empower individuals and enable them to fully participate in health decisions informed by eHealth resources" [23]. Conversely, the extension of digital resources to the health domain in the form of eHealth literacy can also create new gaps between health consumers [14,28]. The digital divide between the haves and have-nots [29] appears to be closing in developed economies in terms of access to the medium; nevertheless, eHealth literacy hinges not on the digital divide but rather on the knowledge gap [30], thus lending support to the hypothesis that information technology is creating a new social inequality, rather than leveling out social discrepancies [28,31].

New inequalities may surface with use of the Internet. Although most people still prefer to receive health information verbally through face-to-face contact with practitioners [12], those with better digital and health knowledge can be expected to consume more information [28] in various forms (whether written or not), and even more so when information is written. Extensive use of digital resources may be associated also with the ability to employ a greater number of search strategies and with a clearer cognizance of the quality of, potential gaps in, and inaccuracies in the obtained information [32]. Finally, the outcomes and benefits of using the Internet for health purposes may extend the traditional outcomes of health literacy [33-36] by providing new areas of physician–patient interaction [37] and self-care.

Present Study and Hypotheses

The present study focused on eHealth literacy and related it to the process and outcome of the information search. First, we examined the structure of the eHealth literacy concept to determine whether the 1-factor structure of the concept, as described by Norman and Skinner, is replicated also in the current study, which used a sample from another culture (Israel) and expanded the age range (compared with the adolescent sample of the original Canadian study [27]). Next, we examined the associations between eHealth literacy and issues related to the digital divide: the factors considered include sociodemographic characteristics, Internet access and digital literacy, the processes involved in information consumption, and the outcomes of using the Internet for health information.

As the sample used by Norman and Skinner was age restricted, we posited no hypothesis about whether our age-expanded sample would replicate the structure of the eHealth scale in the Canadian sample. Following the literature on the digital divide and the digital divide index (DIDIX) in an Israeli sample [15], we hypothesized that people with higher eHealth literacy would be younger and of higher socioeconomic status, would have more access to digital resources, and would exhibit a higher degree of digital literacy than those with lower eHealth literacy.

Following the digital divide literature, we hypothesized the following in regard to the domain of information consumption. People with high eHealth literacy, compared with people low in eHealth literacy, would (1) use more sources of information (magazines, books, television and radio, and interpersonal resources), (2) use a variety of search strategies in addition to googling, (3) judge the information on the Internet more critically and would use more criteria for evaluating health information, and (4) experience more outcomes and in a higher valence as a consequence of using the Internet.

We did not hypothesize about a relationship of eHealth literacy with gender due to shifting findings. Losh [16] and Ono and Zavodny [38] found that gender inequalities in Internet access and usage either diminished, disappeared, or became very specific and context dependent; for example, findings indicate that gender differences shifted to other dimensions such as autonomy of use, experience, skill, and types of uses [17,39,40]. Additionally, findings in an Israeli sample indicate that the DIDIX was lower along gender lines than in any of the other characteristics studied (such as education, income, and age) [15]. We purposely did not present a hypothesis regarding eHealth literacy and health status due to conflicting findings in the literature (eg, Fox [41] vs Bundorf et al [42] on chronic illness; and Goldner [43] vs Wangberg et al [44] on perceived health).

Methods

Data Collection and Sample Characteristics

The current study was conducted as part of a larger study examining gaps between users' needs, proficiencies, and usage processes, on the one hand, and eHealth resources in terms of

quality, language, and assumptions regarding users, on the other hand. Data analyzed in this study were collected from a nationally representative random-digital-dial telephone household survey of the Israeli adult population (18 years of age and older) conducted from May to August 2008 (landlines only). At the time the study was conducted only 7.1% of the Israeli population owned only a mobile phone. Thus, we expected the landline sampling frame to adequately represent the Israeli adult population.

The sampling procedure through which the random digital dialing worked began by dividing statistical areas into 4 layers according to (1) population groups (Jews, Arabs, and mixed localities), (2) 7 geographical districts, (3) different sizes of settlements (big cities to small towns and villages), and (4) socioeconomic status index based on the Israeli Central Bureau of Statistics classification.

Sampling employed a dual-frame design, incorporating two selection stages without stratification in either frame. The larger frame was designed to provide national coverage of the eligible population. Calls were placed to 4286 residential households to identify 2201 eligible potential respondents who use the Internet. Of these respondents, 1289 used the Internet for health purposes. The interviews were conducted by professional interviewers who went through a special training session to familiarize them with the questionnaire's terminology. The interviewers conducted the telephone survey using computer-assisted telephone interviewing software.

A comparison with the Israeli census data indicates that the survey sample in the current study is representative of the Israeli population in terms of sex and age distributions [45]). The survey sample was further controlled to correspond to regional population distribution (see Multimedia Appendix 1).

Measurements

eHealth literacy (perceived) was examined using Norman and Skinner's eHEALS scale [27]. The original scale was composed of 8 questions. Since 2 of the items of the original scale were tapped in our survey in a more detailed fashion, we used only 6 questions. The final scale in our study included 6 items, which met a satisfactory internal consistency criterion (alpha = .86). A 5-point response scale was used (from strongly agree to disagree). Respondents were assigned into two groups on the basis of the mean score they obtained for this 6-item scale of perceived eHealth literacy. The mean score on the scale was 3.34 (SD 0.88). We used the median score of the scale (median 3.4) to create two groups: those with a high mean eHealth literacy score (median ≥3.4); and those with a low mean eHealth literacy score (median ≤3.39).

Internet access was measured by asking participants whether they used the Internet in any of 5 locations at least once a month. A list with 5 locations was presented and participants indicated whether the option applied to them (at a library/community center, friend's, neighbor's, Internet cafe, or school/university).

Digital literacy was tapped by asking for the frequency of engaging in 6 activities (visiting blogs, participating in discussion forums, playing games, downloading or listening to music, downloading software, or emailing with friends). A 5-point frequency response scale was used (from very often to never). Additionally, the user's perceived general Internet skills were tapped (not at all skilled, not very skilled, fairly skilled, very skilled, or expert) [22,46]. The total mean score for digital literacy was computed for each participant (alpha = .75).

Health information *sources* were examined by asking "How often do you get health information from the following sources?" A list of 6 health sources was presented and participants responded to each source. Apart from the Internet as a source of health information, the list included radio or television; newspapers/magazines or books; information obtained from a pharmacist, nurse, or physician; and information obtained from family members or friends. A 5-point response scale was used (from very often to never). A total mean score for the health information channels was computed for each participant (alpha = .64).

Health information *content* on the Internet was examined by asking "How often do you search the Internet for information related to the following domains and actions?" A list of 8 domains and actions related to health information was presented, and participants responded to each domain. The list included seeking information about physicians; institutions that provide health services (hospitals, community clinics, pharmacies, etc); potential treatments (procedures and drugs); and social support. A 5-point response scale was used (from several times a week to never). The total mean score for the seeking health information on the Internet was computed for each participant (alpha = .80).

Search strategies employed to obtain digital health information were examined by asking "In order to find health information on the Internet you usually do the following." A list of 5 common search actions was presented: use a site that my physician recommended; follow links that appear on websites; ask questions in forums; use my Favorites list; and use a site that a friend recommended. A 5-point frequency response scale was used (from always to never). The total mean score for the health information search strategy was computed for each participant (alpha = .64).

*Evaluation criteri*a were examined using Barnes and colleagues' [32] scale. Participants were asked how important the 5 following criteria were in judging a website: the purpose of the site is clearly stated and the information is accurate; it has a reliable source; a contact is available for questions/comments/help; retrieval is easy and can be done in a timely manner; and the scope of information suits my needs. A 5-point response scale was used (from very important to don't know). The total mean score for the health website evaluation criteria was computed for each participant (alpha = .77).

Perceived outcomes of seeking health information on the Internet were examined by asking "Do you agree or disagree that seeking health information on the Internet...?" A list of 9 outcomes, adapted from Baker et al [47], was presented: improved your ability to manage your health needs; enabled you to ask your physician questions resulting from the information you acquired on the Internet; enabled you to show your physician the information that you retrieved; raised your sense of power in your encounter with the physician; improved

your understanding of the symptoms, conditions, or treatments in which you were interested; updated your knowledge in health innovations; led you to take independent steps (such as seeing a specialist, or changing an exercise regimen or eating habits); enabled you to think about alternative treatment options; and made you more aware of patients' insurance rights (all Israeli citizens have a health insurance). A 5-point response scale was used (from strongly agree to disagree). The total mean score for each participant's outcome perception was computed (alpha = .87).

Sociodemographic information included sex, age, levels of obtained education, religiosity, perceived health condition, and chronic diseases.

Perceived health tapped respondents' self-rated health, as compared with other people their age and gender. Respondents indicated whether they were about the average, somewhat above the average, much above the average, somewhat below the average, or much below the average.

The existence of *chronic diseases* was assessed by asking respondents whether they had any chronic diseases.

Data Analysis

We first conducted principal components analyses on the eHEALS. Second, to indicate that our measures were separate factors, we conducted two sets of confirmatory factor analysis on all major variables, using the CALIS procedure of SAS version 9.2 (SAS Institute, Cary, NC, USA). Lastly, we used analysis of variance and χ^2 to compare differences between groups, and Pearson correlation to examine relations between items or between indices.

Results

Exploratory Factor Analysis of the eHealth Literacy Scale

As the current study used only 6 of the 8 original eHEALS items, we conducted an exploratory factor analysis on these items. Principal components analysis produced a single-factor solution (eigenvalue = 3.551, 59% of the variance explained). Factor loadings ranged from .62 to .84 among the 6 items. Internal consistency reliability was analyzed on the 6 items, producing a coefficient alpha of .86, where item–scale correlations ranged from r = .50 to .73. These results are quite similar to those of Norman and Skinner [27], where the single-factor solution explained 56% of the variance, the internal consistency reliability was alpha = .88, and the item–scale correlations ranged from r = .51 to .76.

Confirmatory Factor Analysis of the Research Scales

We calculated 2-model fit analyses to insure the assumption that each scale is independent of the other scales. In the first set of analyses, we used confirmatory factor analysis to test the structure of 4 scales: eHealth literacy, outcomes perception, digital literacy, and Internet access. The results confirmed that the 4 scales are independent of each other, and the 6-item scale used from Norman and Skinner's [27] eHEALS is considered an independent scale.

In the second set of confirmatory factor analyses we tested the independence of an additional 5 scales in our study: health information sources, health information content, motivations for information search, search strategy, and evaluation criteria. The results confirmed that the 5 scales are considered independent of each other (see Multimedia Appendix 2).

Sociodemographic Characteristics of the eHealth Literacy Groups

We examined the characteristics of the high and low eHealth literacy groups focusing first on the demographic variables of gender, age, and socioeconomic status. The high and low eHealth literacy groups did not differ in gender. There were 321 (50.7%) and 325 women (49.8%) in the high (n = 633) and low groups (n = 653), respectively (χ^2_1 = 0.11, P > .05). Likewise, the eHealth literacy score of men (mean 3.35, SD 0.89) and women (mean 3.31, SD 0.88) did not differ significantly ($F_{1,1284}$ = 0.94, P = .332). However, the high eHealth literacy group was significantly younger ($F_{1,1284}$ = 35.56, P < .000; mean 38.87, SD 14.40, years) than the low eHealth literacy group (mean 44.12, SD 17.00, years). The socioeconomic status of the high eHealth literacy group was also significantly higher than that of the low eHealth literacy group, as measured by education (mean score, on a 7-point scale, 3.99, SD 1.32 and mean 3.82, SD 1.33, for the high and low eHealth literacy groups, respectively, $F_{1,1274}$ = 5.43, P < .02). There were 264 (41.9%) and 228 (35.2%) respondents with academic degrees, respectively, in the high (n = 630) and low (n = 647) eHealth literacy groups.

The health status of the eHealth groups was significantly different between the eHealth literacy groups. Respondents who reported that they were chronically ill had a significantly lower eHealth literacy score ($F_{1,1270}$ = 8.87, P < .003; mean 3.19, SD 0.95) than respondents with no reported chronic illnesses (mean 3.37, SD 0.85). In addition, 164 respondents (25.3%) in the lower eHealth literacy group (n = 648) reported having a chronic illness, as compared with only 117 (18.8%) respondents in the higher eHealth literacy group (n = 624). The health status difference on eHealth literacy was independent of age: an analysis of variance on eHealth literacy revealed an insignificant interaction effect of health status and age ($F_{3,1262}$ = 0.695, P =.44).

Health status was also examined in terms of perceived health. There was no significant difference between the high and low eHealth literacy groups in perceived health ($F_{1,1276}$ = 0.432, P =.511). The high and low eHealth literacy groups reported similar self-rated health (mean 3.25, SD 0. 74 and mean 3.22, SD 0.68, for the high and low eHealth literacy groups, respectively).

Internet Access and Digital Literacy

eHealth literacy emerged as related to digital access and literacy. Respondents in the high eHealth literacy group had significantly more access to computers and used the Internet more frequently than did the low eHealth literacy group ($F_{1,1281}$ = 26.47, P < .001): the mean Internet accessibility score of the high eHealth literacy group was 6.19, as compared with a score of 5.86 among

the low eHealth literacy group. Furthermore, the digital literacy reported by the high eHealth literacy group was significantly higher than that reported by the low eHealth literacy group: 2.67 and 2.24, for the high and low eHealth literacy groups, respectively ($F_{1,1280}$ = 88.34, P < .001).

Information Consumption: Health Information Sources, Health Information Content on the Internet, Health Website Evaluation Criteria

eHealth literacy is a marker for consuming more information, as displayed in Table 1. Overall, respondents in the high eHealth literacy group used significantly more information sources ($F_{1,1280}$ = 11.01, P < .001) than did the low eHealth literacy group. Looking at individual items in Table 1, there is a significant difference between the two groups in their use of written material such as books, newspapers, magazines, and the Internet; there is no statistically significant difference between the two eHealth literacy groups in their use of live information from radio and television, a pharmacist, a nurse, or a physician (all P > .05).

Table 1. Scores for low and high eHealth literacy groups' consumption of information on the Internet

Variable	Low		High		*F*	*P* value
	Mean	SD	Mean	SD		
Information source (index)	2.62	0.65	2.75	0.73	11.01	<.001
Books	2.23	1.16	2.59	1.34	26.30	<.001
Newspapers and magazines	2.64	1.14	2.87	1.22	12.40	<.001
Internet	3.26	0.92	3.81	0.93	112.78	<.001
Radio and television	2.81	1.13	2.79	1.23	0.68	.794
Pharmacist	2.21	1.25	2.15	1.26	0.86	.355
Nurse or physician	3.41	1.25	3.54	1.26	3.55	.06
Search quantity/variety (index)	1.75	0.59	2.05	0.73	66.28	<.001
Variety of search strategies (index)	2.16	0.74	2.59	0.90	87.08	<.001
Information evaluation (index)	4.29	0.68	4.53	0.50	52.21	<.001

Respondents in the high eHealth literacy group searched for significantly more content on the Internet ($F_{1,1280}$ = 66.28, P < .001) than did the low eHealth literacy group, irrespective of the type of health content: social (eg, social support groups), service-related (eg, availability of services, or information on physicians, hospitals, and pharmacies), and therapy-related content (eg, health status, procedures, and medication).

The use of the Internet was different in terms of the search strategies employed by each of the two eHealth literacy groups. As can be seen in Table 1, those high in eHealth literacy used every strategy significantly more often than those low in eHealth literacy ($F_{1,1280}$ = 87.08, P < .000). For example, they followed links, asked questions on Internet forums, followed recommendations of their friends and physicians, and used their Favorites list significantly more often than those low in eHealth literacy.

In addition, the use of the Internet by participants who scored high on the eHealth literacy scale was marked by significantly more scrutiny, caution, and evaluation of the information they retrieve ($F_{1,1280}$ = 52.21, P < .000). Thus, for example, they looked for a contact address, wondered about the reliability of the source and the accuracy of information, and formed an opinion about the accessibility and availability of the information on the particular site they encountered. Information evaluation is related to eHealth literacy (r = .26, P < .000), but as the size of the correlation indicates, it is not synonymous with eHealth literacy.

Outcomes of Information Search

Finally, those highly eHealth literate gained significantly more from their information search than did the low eHealth literacy group (3.40 and 2.76, for high and low eHealth literacy groups, respectively; $F_{1,1280}$ = 177.76, P < .001). The results are displayed in Table 2. Cognitively, people in the high eHealth literacy group reported gaining a better understanding of their health status, symptoms, and optional treatments (see items in Table 2). They also benefited more instrumentally: the information search improved their ability to self-manage their health care needs, affected their health behaviors, and allowed them a better use of their health insurance. The benefits extended also to their interaction with the treating physician: they asked the physician significantly more questions than they would have without the digital information search, presented the physician with the information they retrieved, and felt significantly better positioned vis-à-vis the physician than did the low eHealth literacy group.

Table 2. Scores for low and high eHealth literacy groups in outcomes of information search

Variable	Low		High		*F*	*P* value
	Mean	SD	Mean	SD		
Outcomes (index)	2.76	0.88	3.40	0.83	177.76	<.001
Understanding of symptoms, conditions, treatment	3.30	1.20	3.95	0.96	115.56	<.001
Update in health innovations	3.01	1.24	3.71	1.16	108.04	<.001
Self-managing health	2.37	1.24	3.13	1.34	87.39	<.001
Affected health behaviors	2.75	1.25	3.41	1.25	87.39	<.001
Use of insurance	2.23	1.33	2.77	1.43	45.95	<.001
Asking physician questions	3.17	1.28	3.73	1.18	63.51	<.001
Consulting physician on information retrieved	2.90	1.32	3.54	1.24	81.85	<.001
Power position with physician	2.55	1.32	3.22	1.31	83.06	<.001

Discussion

Principal Results

The present study has demonstrated the utility of the eHealth literacy concept. Though we used only 6 of the original 8 items of the eHEALS, the 1-factor structure of the construct emerged also in the current Israeli sample, indicating that the concept of eHealth literacy is applicable to another culture and age groups.

The main contributions of this study, however, lie in demonstrating the relation between eHealth literacy and (1) the background attributes of the respondents, (2) patterns of information consumption, and (3) outcomes of the information search. In almost all 3 criteria, findings showed that the degree of eHealth literacy skills extended the digital divide into the health domain. The implication of these findings is that low eHealth-literate people would be limited in their use of the resources available on the Internet.

We hypothesized that respondents higher in eHealth literacy would be younger and more educated. This hypothesis was supported. We put forward no hypotheses regarding gender and health status (measured by perceived health and chronic illness), and the findings indicate no gender or perceived health differences; however, we noted a significant difference in health status: the chronically ill were lower in eHealth literacy. Hypotheses regarding greater digital access and higher digital literacy among those with high eHealth literacy were supported, as well as all of the hypotheses regarding information consumption: using more information sources, conducting more frequent and more varied searches, employing more search strategies, and evaluating the output. Moreover, as hypothesized, respondents higher in eHealth literacy used the information gained more than did respondents low in eHealth literacy. Indeed, for those who can realize the potential and possibilities, the Internet is a means of sustaining health, whether by providing information, linking to peers and professionals, or supporting self-management of health and illness [48]. However, the use of the Internet in the health domain is related to social inequality [49]: health information was already identified as capital-enhancing activity (vs recreational activity) [17,49], and the present findings indicate who among the connected benefits the most and in what ways. We found that differences in traditional variables of the digital divide literature (age, education, health status, digital access, and literacy) were associated with eHealth literacy, and we recorded new hypothesized differences in information consumption and outcomes gained from the use of the Internet. As theoreticians surmised [50,51], crossing the initial connectivity divide left numerous differences between people in how they incorporated the Internet into their lives.

Comparison With Prior Work

The findings regarding the relationship between background characteristics and eHealth literacy are similar to findings obtained in studies documenting digital access and digital literacy disparities [13,20,22,30]. As expected, the eHealth literacy groups differed significantly in terms of education and age, duplicating differences found between those who have access to computers and the Internet and know how to use them [10,12,14] and those who do not know how to use computers and the Internet. Gender did not differ between our eHealth literacy groups, a finding congruent with others' [16,38]. It demonstrates the conclusion that gender differences are highly contextualized. It is also congruent with Mizrachi et al [15], who found that the DIDIX along gender lines in an Israeli sample was the lowest among other characteristics (such as education, income, and age). Finally, and as hypothesized, the eHealth literacy groups were significantly different in digital access and digital literacy. Digital access avails the information search and digital literacy is conceptually a facet of eHealth literacy [23].

Health status was measured in our study in two ways: perceived health (ie, self-rated health) and reported chronic illnesses. Perceived health did not vary with eHealth literacy but the presence of chronic illnesses did, such that chronically ill respondents had lower eHealth literacy scores. The inconsistency between the two measures is acceptable, as the two assess different concepts; for example, a person may have a diagnosis of hypertension, but she may also feel healthier than most of her age group. Still, our findings on the two measures replicate some previous work [41,43] and contrast with others [42,44], and no clear picture yet emerges. It could be that the relationship between health status and information search is dependent on

the health system. Indeed, Bundorf et al [42] suggest that searching the Internet for health depends on costs and benefits, such as paying out-of-pocket and opportunity costs; these vary across health systems and may explain the seemingly contradictory findings. The present findings of lower eHealth literacy among the chronically ill call for an empowering intervention on the part of service providers (for example, health maintenance organizations). The chronically ill population constitutes a highly distinguished group for service providers, and its eHealth literacy should be targeted for improvement. Planned learning experiences to improve their literacy in searching for, locating, evaluating, and using eHealth information is called for. As Hargittai [17] points out, "achieving a knowledgeable Internet citizenry is unlikely to be resolved through a solely technical approach that focuses only on infrastructure without any consideration of the social processes and institutions in which people's Internet uses are embedded."

The findings on consuming information—ways of searching, frequency of searching for various contents, evaluating the information, benefiting from the information—all demonstrate inequality among those high and low in eHealth literacy. All of these variables focus on utilization rather than on mere accessibility [52]; these variables exemplify new differentiated usage patterns among the connected that have the potential to contribute to social inequality [17,50,51,53,54]. These findings are more in line with the strong hypothesis than with the weak hypothesis of the digital divide. The strong hypothesis posits that "the emergence of the information society will create new social cleavages and strengthen old ones" [28], whereas the weak hypothesis claims that the new technology will level out old differences, admittedly after witnessing a temporary gap during the dissemination of the new technology. Still, it is possible that the weak hypothesis is not altogether amiss. We may be in the midst of a change, as exemplified in gender, which was in the past related to digital access and literacy and turned out to be unrelated to eHealth literacy in this study.

It is important to note that the fit or full utilization of technology depends not only on users and their characteristics but also on the technology itself. Technology may have different manifestations and affordances [55], and it could be tailored to fit different users [56,57]. The literature on informed choice demonstrates [58] that there are many ways of presenting information and choices, and it exhibits genuine efforts and achievements in presenting complex medical information to laypeople. Similar efforts are called for in the content, design, and ease of use of health information on the Internet, so that people low in eHealth literacy may make fuller use of the digital promise. The realization of such a promise may also call for setting standards and accreditation, and alerting consumers to the latter.

The results of the confirmatory factor analysis indicate that the concept of eHealth literacy is independent of related variables. Even the variable of *health website evaluation criteria*, which seems to measure media literacy (in this case, Web literacy), was found to be only moderately related to eHealth literacy (r = .26, P < .000). This may indicate that, as Norman and Skinner [23] point out, media literacy is only one aspect of eHealth literacy.

Limitations

Our findings are hampered by 5 major limitations. First, we used only 6 items from the original eHEALS [27]. Even though the 1-factor structure of the concept, the internal reliability, and the item–scale correlations were highly similar to the original scale, the scales are not psychometrically equivalent. Second, the cross-sectional design of the study precludes causal conclusions and allows us to draw conclusions only regarding correlated relationships. For example, we can assume only that education is associated with eHealth literacy and not that it affects eHealth literacy. Third, we did not measure actual eHealth literacy but rather perceived efficacy of searching and using health information on the Internet. This limitation calls for an amendment. Measures of actual eHealth literacy are required. Indeed, there are measures for actual digital literacy [39,59] and there are measures for health literacy [60]. The two may serve as an inspiration for a measure that taps actual eHealth literacy. The fourth limitation is related to the third: our findings are based on self-reports and not actual performance or record of Internet use. More studies that measure actual use and skill [48] are needed. Finally, the fifth limitation is related to the landline sampling frame. This sampling frame excluded 7.1% of the population who owned only mobile phones at the time that the survey was conducted. These people may be younger and may have been more digitally and eHealth literate.

To conclude, the present study documented differences between respondents high and low in eHealth literacy in terms of background attributes, information consumption, and outcomes of the information search. The findings are mostly interpreted in line with the digital divide literature, replicating previously demonstrated relationships to background variables (demographics, digital access, and literacy), and identifying and documenting new cleavages (information consumption and perceived outcomes). The need to both educate at-risk and needy groups [61] (eg, chronically ill) and design technology in a mode befitting more consumers emerges. Addressing those needs may not diminish the digital divide altogether, but it may ameliorate its consequences by bringing more people into the have group.

Acknowledgments

The study was supported by the Israel National Institute for Health Policy and Health Services Research.

The authors wish to thank the JMIR reviewers for their valuable suggestions for improving the article.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristic of the participant's sex and age compared with the population.

[PDF File (Adobe PDF File), 57KB - jmir_v14i1e19_app1.pdf]

Multimedia Appendix 2

Scales independence calculation through confirmatory factor analysis (CFA).

[PDF File (Adobe PDF File), 23KB - jmir_v14i1e19_app2.pdf]

References

1. Oh H, Rizo C, Enkin M, Jadad A. What is eHealth (3): a systematic review of published definitions. J Med Internet Res 2005;7(1):e1 [FREE Full text] [doi: 10.2196/jmir.7.1.e1] [Medline: 15829471]
2. Eng TR. The eHealth Landscape: A Terrain Map of Emerging Information and Communication Technologies in Health and Health Care. Princeton, NJ: Robert Wood Johnson Foundation; 2001.
3. Nutbeam D. World Health Organization. 1998. Health Promotion Glossary URL: http://www.who.int/healthpromotion/about/HPR%2520Glossary%25201998.pdf [accessed 2012-01-26] [WebCite Cache ID 64z2Lv6Km]
4. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Promot Int 2000;15(3):259-267. [doi: 10.1093/heapro/15.3.259]
5. Hasnain-Wynia R, Wolf MS. Promoting health care equity: is health literacy a missing link? Health Serv Res 2010 Aug;45(4):897-903 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01134.x] [Medline: 20646073]
6. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624 [FREE Full text] [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
7. Kontos EZ, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. J Med Internet Res 2007;9(4):e31 [FREE Full text] [doi: 10.2196/jmir.9.4.e31] [Medline: 17951215]
8. McNeill LH, Puleo E, Bennett GG, Emmons KM. Exploring social contextual correlates of computer ownership and frequency of use among urban, low-income, public housing adult residents. J Med Internet Res 2007;9(4):e35 [FREE Full text] [doi: 10.2196/jmir.9.4.e35] [Medline: 18093903]
9. Kuntalp M, Akar O. A simple and low-cost Internet-based teleconsultation system that could effectively solve the health care access problems in underserved areas of developing countries. Comput Methods Programs Biomed 2004 Aug;75(2):117-126. [doi: 10.1016/j.cmpb.2003.11.004] [Medline: 15212854]
10. Fleming J. Health information on the Internet. J R Soc Promot Health 2003 Mar;123(1):10-11. [Medline: 12722578]
11. Mead N, Varnam R, Rogers A, Roland M. What predicts patients' interest in the Internet as a health resource in primary care in England? J Health Serv Res Policy 2003 Jan;8(1):33-39. [doi: 10.1258/135581903360468209] [Medline: 12683432]
12. Välimäki M, Nenonen H, Koivunen M, Suhonen R. Patients' perceptions of Internet usage and their opportunity to obtain health information. Med Inform Internet Med 2007 Dec;32(4):305-314. [doi: 10.1080/14639230701819792] [Medline: 18072007]
13. Buente W, Robbin A. Trends in Internet information behavior, 2000-2004. J Am Soc Inf Sci Technol 2008;59(11):1743-1760. [doi: 10.1002/asi.20883]
14. Miller EA, West DM. Characteristics associated with use of public and private web sites as sources of health care information: results from a national survey. Med Care 2007 Mar;45(3):245-251. [doi: 10.1097/01.mlr.0000244509.60556.49] [Medline: 17304082]
15. Mizrachi Y, Bar N, Katsernov I, Oron N. e-Readiness and Digital Divide survey, Israel 2005 [Hebrew]. Jerusalem: Ministry of Finance; 2005. URL: http://www.maor.gov.il/Maor/Docs/HE/DigitalSurvey/digital3.pdf [WebCite Cache ID 64ypRnMNi]
16. Losh SC. Gender, educational, and occupational digital gaps 1983-2002. Social Sci Comput Rev 2004;22(2):152-166. [doi: 10.1177/0894439303262557]
17. Hargittai E. Digital na(t)ives? Variation in Internet skills and use among members of the. Sociol Inq 2010;80(1):92-113. [doi: 10.1111/j.1475-682X.2009.00317.x]
18. Barzilai-Nahon K. Gaps and bits: conceptualizing measurements for digital divide. Inform Soc 2006;22(5):269-278. [doi: 10.1080/01972240600903953]
19. Skinner H, Biscope S, Poland B. Quality of internet access: barrier behind internet use statistics. Soc Sci Med 2003 Sep;57(5):875-880. [Medline: 12850112]

20. Zillien N, Hargittai E. Digital distinction: status-specific types of Internet usage. Soc Sci Q 2009;90(2):274-291. [doi: 10.1111/j.1540-6237.2009.00617.x]
21. Vehovar V, Sicherl P, Hüsing T, Dolnicar V. Methodological challenges of digital divide measurements. Inform Soc 2006;22(5):279. [doi: 10.1080/01972240600904076]
22. Hargittai E. Second-level digital divide: differences in people's online skills. First Monday 2002;7(4).
23. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
24. Miller EA, West DM. Where's the revolution? Digital technology and health care in the internet age. J Health Polit Policy Law 2009 Apr;34(2):261-284. [doi: 10.1215/03616878-2008-046] [Medline: 19276318]
25. 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]
26. Pálsdóttir Á. Seeking information about health and lifestyle on the Internet. Inform Res 2009;14(1) [FREE Full text]
27. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
28. Sassi S. Cultural differentiation or social segregation? Four approaches to the digital divide. New Media & Society 2005;7(5):684. [doi: 10.1177/1461444805056012]
29. Fox S. Digital divisions. Washington, DC: Pew Internet & American Life Project; 2005 Oct 5. URL: http://www.pewinternet.org/~/media//Files/Reports/2005/PIP_Digital_Divisions_Oct_5_2005.pdf.pdf [accessed 2012-01-26] [WebCite Cache ID 64ykcVepk]
30. Viswanath K, Finnegan JR. The knowledge gap hypothesis: twenty-five years later. In: Burleson BR, editor. Communication Yearbook 19. Thousand Oaks, CA: Sage; 1996:187.
31. Stevenson S. Digital divide: a discursive move away from the real inequities. Inform Soc 2009;25(1):1-22. [doi: 10.1080/01972240802587539]
32. Barnes MD, Penrod C, Neiger BL, Merrill RM, Thackeray R, Eggett DL, et al. Measuring the relevance of evaluation criteria among health information seekers on the Internet. J Health Psychol 2003 Jan;8(1):71-82. [doi: 10.1177/1359105303008001436] [Medline: 22113902]
33. Baker DW, Parker RM, Williams MV, Pitkin K, Parikh NS, Coates W, et al. The health care experience of patients with low literacy. Arch Fam Med 1996 Jun;5(6):329-334 [FREE Full text] [Medline: 8640322]
34. Montalto NJ, Spiegler GE. Functional health literacy in adults in a rural community health center. W V Med J 2001;97(2):111-114. [Medline: 11392190]
35. Nurss JR, el-Kebbi IM, Gallina DL, Ziemer DC, Musey VC, Lewis S, et al. Diabetes in urban African Americans: functional health literacy of municipal hospital outpatients with diabetes. Diabetes Educ 1997;23(5):563-568. [Medline: 9355373]
36. Williams MV, Parker RM, Baker DW, Parikh NS, Pitkin K, Coates WC, et al. Inadequate functional health literacy among patients at two public hospitals. JAMA 1995 Dec 6;274(21):1677-1682. [Medline: 7474271]
37. Taub SJ. The Internet's role in patient/physician interaction: bringing our understanding in line with online reality. Compr Ophthalmol Update 2006;7(1):25-30. [Medline: 16630413]
38. Ono H, Zavodny M. Gender and the Internet. Soc Sci Q 2003;84(1):111-121. [doi: 10.1111/1540-6237.t01-1-8401007]
39. Hargittai E. Classifying and coding online actions. Soc Sci Comput Rev 2004;22(2):210-227. [doi: 10.1177/0894439303262560]
40. Van Deursen AJAM, van Dijk JAGM, Peters O. Rethinking Internet skills: the contribution of gender, age, education, Internet experience, and hours online to medium- and content-related Internet skills. Poetics 2011;39(2):125-144. [doi: 10.1016/j.poetic.2011.02.001]
41. Fox S. E-patients with a disability or chronic disease. Washington, DC: Pew Internet & American Life Project; 2007 Oct 8. URL: http://www.pewinternet.org/~/media/Files/Reports/2007/EPatients_Chronic_Conditions_2007.pdf.pdf [accessed 2012-01-26] [WebCite Cache ID 64ym00GQq]
42. Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? Health Serv Res 2006 Jun;41(3 Pt 1):819-836. [doi: 10.1111/j.1475-6773.2006.00510.x] [Medline: 16704514]
43. Goldner M. Using the Internet and email for health purposes: the impact of health status. Soc Sci Q 2006;87(3):690-710. [doi: 10.1111/j.1540-6237.2006.00404.x]
44. Wangberg SC, Andreassen HK, Prokosch HU, Santana SM, Sørensen T, Chronaki CE. Relations between Internet use, socio-economic status (SES), social support and subjective health. Health Promot Int 2008 Mar;23(1):70-77 [FREE Full text] [doi: 10.1093/heapro/dam039] [Medline: 18083686]
45. Israel Central Bureau of Statistics. Statistical Abstract of Israel. 2007. Table 2.10: Population, by Population Group, Religion, Age and Sex, District and Sub-district URL: http://www1.cbs.gov.il/reader/shnaton/templ_shnaton_e.html?num_tab=st02_10x&CYear=2007 [accessed 2012-01-17] [WebCite Cache ID 64yw36i8a]
46. Hargittai E. Survey measures of Web-oriented digital literacy. Soc Sci Comput Rev 2005;23(3):371-379. [doi: 10.1177/0894439305275911]

47. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406 [FREE Full text] [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
48. 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] [Medline: 20205015]
49. Howard PN, Rainie L, Jones S. Days and nights on the Internet: the impact of a diffusing technology. Am Behav Sci 2001 Sep 1;45(3):383-404. [doi: 10.1177/0002764201045003003]
50. van Dijk JAGM. The Deepening Divide: Inequality in the Information Society. Thousand Oaks, CA: Sage Publications, Inc; 2005.
51. DiMaggio P, Hargittai E, Celeste C, Shafer S. Digital inequality: from unequal access to differentiated use. In: Neckerman KM, editor. Social Inequality. New York, NY: Russell Sage Foundation; 2004:355-400.
52. DiMaggio P, Hargittai E, Neuman WR, Robinson JP. Social implications of the Internet. Annu Rev Sociol 2001;27:307-336. [doi: 10.1146/annurev.soc.27.1.307]
53. Hargittai E. The digital reproduction of inequality. In: Grusky DB, editor. Social Stratification: Class, Race, and Gender in Sociological Perspective, 3rd edition. Boulder, CO: Westview Press; 2008:936-944.
54. Selwyn N. Reconsidering political and popular understandings of the digital divide. New Media Soc 2004;6(3):341-362. [doi: 10.1177/1461444804042519]
55. Gibson EJ. Perceiving the Affordances: A Portrait of Two Psychologists. Mahwah, NJ: Erlbaum; 2002.
56. Xie B. Effects of an eHealth literacy intervention for older adults. J Med Internet Res 2011;13(4):e90 [FREE Full text] [doi: 10.2196/jmir.1880] [Medline: 22052161]
57. Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. J Med Internet Res 2011;13(3):e75 [FREE Full text] [doi: 10.2196/jmir.1697] [Medline: 21960017]
58. Woolf SH, Chan EC, Harris R, Sheridan SL, Braddock CH, Kaplan RM, et al. Promoting informed choice: transforming health care to dispense knowledge for decision making. Ann Intern Med 2005 Aug 16;143(4):293-300. [Medline: 16103473]
59. Hargittai E, Shafer S. Differences in actual and perceived online skills: the role of gender. Soc Sci Q 2006;87(2):432-448. [doi: 10.1111/j.1540-6237.2006.00389.x]
60. McCray AT. Promoting health literacy. J Am Med Inform Assoc 2005;12(2):152-163 [FREE Full text] [doi: 10.1197/jamia.M1687] [Medline: 15561782]
61. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. J Med Internet Res 2011;13(4):e94 [FREE Full text] [doi: 10.2196/jmir.1750] [Medline: 22094891]

Abbreviations

DIDIX: digital divide index
eHEALS: eHealth literacy scale

Edited by G Eysenbach; submitted 19.07.10; peer-reviewed by C Norman, I O'Boyle; comments to author 13.11.10; revised version received 20.10.11; accepted 18.11.11; published 27.01.12.

Please cite as:
Neter E, Brainin E
eHealth Literacy: Extending the Digital Divide to the Realm of Health Information
J Med Internet Res 2012;14(1):e19
URL: http://www.jmir.org/2012/1/e19/
doi:10.2196/jmir.1619
PMID:22357448

Original Paper

Development and Implementation of a Web-Enabled 3D Consultation Tool for Breast Augmentation Surgery Based on 3D-Image Reconstruction of 2D Pictures

Pablo de Heras Ciechomski[1*], PhD; Mihai Constantinescu[2*], MD; Jaime Garcia[3], PhD; Radu Olariu[2], MD; Irving Dindoyal[1], PhD; Serge Le Huu[4], MD; Mauricio Reyes[1*], PhD

[1]Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland

[2]Department of Plastic, Reconstructive, and Aesthetic Surgery, University Hospital, Inselspital, University of Bern, Bern, Switzerland

[3]Crisalix SA, Lausanne, Switzerland

[4]La Clinic, Montreux, Switzerland

[*]these authors contributed equally

Corresponding Author:
Mauricio Reyes, PhD
Institute for Surgical Technology and Biomechanics
University of Bern
Staufacherstrasse 78
Bern, 3014
Switzerland
Phone: 41 316315950
Fax: 41 316315960
Email: mauricio.reyes@istb.unibe.ch

Abstract

Background: Producing a rich, personalized Web-based consultation tool for plastic surgeons and patients is challenging.

Objective: (1) To develop a computer tool that allows individual reconstruction and simulation of 3-dimensional (3D) soft tissue from ordinary digital photos of breasts, (2) to implement a Web-based, worldwide-accessible preoperative surgical planning platform for plastic surgeons, and (3) to validate this tool through a quality control analysis by comparing 3D laser scans of the patients with the 3D reconstructions with this tool from original 2-dimensional (2D) pictures of the same patients.

Methods: The proposed system uses well-established 2D digital photos for reconstruction into a 3D torso, which is then available to the user for interactive planning. The simulation is performed on dedicated servers, accessible via Internet. It allows the surgeon, together with the patient, to previsualize the impact of the proposed breast augmentation directly during the consultation before a surgery is decided upon. We retrospectively conduced a quality control assessment of available anonymized pre- and postoperative 2D digital photographs of patients undergoing breast augmentation procedures. The method presented above was used to reconstruct 3D pictures from 2D digital pictures. We used a laser scanner capable of generating a highly accurate surface model of the patient's anatomy to acquire ground truth data. The quality of the computed 3D reconstructions was compared with the ground truth data used to perform both qualitative and quantitative evaluations.

Results: We evaluated the system on 11 clinical cases for surface reconstructions and 4 clinical cases of postoperative simulations, using laser surface scan technologies showing a mean reconstruction error between 2 and 4 mm and a maximum outlier error of 16 mm. Qualitative and quantitative analyses from plastic surgeons demonstrate the potential of these new emerging technologies.

Conclusions: We tested our tool for 3D, Web-based, patient-specific consultation in the clinical scenario of breast augmentation. This example shows that the current state of development allows for creation of responsive and effective Web-based, 3D medical tools, even with highly complex and time-consuming computation, by off-loading them to a dedicated high-performance data center. The efficient combination of advanced technologies, based on analysis and understanding of human anatomy and physiology, will allow the development of further Web-based reconstruction and predictive interfaces at different scales of the human body. The consultation tool presented herein exemplifies the potential of combining advancements in the core areas of computer science and biomedical engineering with the evolving areas of Web technologies. We are confident that future developments based on a multidisciplinary approach will further pave the way toward personalized Web-enabled medicine.

(J Med Internet Res 2012;14(1):e21) doi:10.2196/jmir.1903

KEYWORDS

Medical informatics computing; computer-assisted surgery; imaging, three-dimensional

Introduction

Since the creation of the World Wide Web in the early 1990s, its use for medical applications has attracted much attention due to the possibilities of centralized storage and the efficient sharing of information. The creation of the picture archiving and communication system and related Web-enabled interfaces for the Internet demonstrates the interest from the medical community in accessing information in a reliable, economical, and convenient way [1,2]. However, despite efforts in computer sciences (eg [3,4]), the processing of medical images is still computationally expensive for real-time use on most personal computers. Continuing efforts are being made toward personalized patient models for the predictive health care of the future [5], leading to new pathways in health care [6].

A field in which Internet capabilities can be used for medical purposes is 3-dimensional (3D) human anatomy. Contrary to Web-enabled medical tools for education purposes, where standard data models are employed, the scenario is more complex when considering confidential patient-specific or personalized medical imaging data from 2-dimensional (2D) pictures. Therefore, the tool presented in this paper was developed and tested in a multidisciplinary effort by a team of experts consisting of surgeons, biomedical engineers, computer graphic specialists, and Web developers and designers.

In breast augmentation surgery, surgeon–patient communication is vital, as the diagnosis, treatment, and outcome are dominated by the patient's subjective assessment of the visual results of the elective surgical procedure. Failure to meet the patient's expectations (augmentation volume, breast projection, etc) can lead to the need for reoperations and ultimately to legal action. It is therefore essential that patients be personally involved in the process of implant selection, supported by a realistic visual representation of their body, the previsualization of the final result. The success of the surgical outcome depends significantly on the choice of implant shape, size, projection, and anatomical placement, and these are key factors in the decision process.

Available computerized 3D anatomical visualization can be divided into the following categories. First, image-morphing techniques are software solutions working exclusively in 2 dimensions, where a patient's photograph might or might not be the basis for the projected postoperative result (eg, PhotoShop, ReShapr [7], PlasticSurgerySimulator [8]). Second, templated and predefined software allows a user to define a set of parameters and to relate them to a predefined model and a predefined outcome (eg, BreastDoctors [9], LoveYourLook [10]). Third, educational software has been one of the areas where 3D Web-based medicine has shown success. For instance, the use of avatars in virtual reality learning environments [11-13] has captured the attention of researchers and medical practitioners due to the dynamic and engaging learning, peer collaboration, and interaction with users around the world [14-18]. However, these do not meet the requirements of individualized patient data analysis. Fourth, 3D scans allow an accurate 3D reconstruction of the patient's specific shape, texture, color, and sizing. Most scanning hardware such as Portrait3D (www.axisthree.com) and VECTRA 3D (www.canfieldsci.com) requires bulky and costly equipment that must be installed at each surgeon's office.

The above-mentioned techniques have inherent limitations for application in the daily clinical work of a surgeon.

We propose a patient-specific system for breast augmentation previsualization using well-established 2D digital photos of the patient's body taken with a digital camera (together with a few extra body measurements for scaling) and transforming them into a 3D, interactive, visual surface representation of the upper torso, which is then available for interactive planning to the user. The simulation is performed on dedicated servers, accessible via Internet. It allows the surgeon, together with the patient, to previsualize the impact of the proposed breast augmentation directly during the consultation before a surgery is decided upon. The hypothesis of this study was thus that the proposed Web-based system would allow previsualization of results, with varying implant size and varying implant locations, acting as a guide for the preoperative planning and decision process.

The 3 goals of this study were thus to (1) develop a computer tool that allows the individual reconstruction and simulation of 3D soft tissue from ordinary digital photos of breasts, (2) implement a Web-based, worldwide-accessible preoperative surgical planning platform for plastic surgeons, and (3) validate this tool through a quality control analysis by comparing 3D laser scans of the patients with the 3D reconstructions made using this tool from original 2D pictures of the same patients.

Methods

The following subsections describe some particular adopted strategies, giving particular emphasis to Web-related components and data management.

Figure 1 illustrates the general pipeline of the developed system. Patients and clinicians are connected through the Internet, enabling a dialogue while being independent of geographical location. Personalized patient information, such as digital photographs and sparse measurements, enable the modeling and simulation framework to create a completely patient-specific clinical scenario.

Figure 1. General overview of the developed system. An Internet-based solution combining advanced technologies enables a realistic, patient-specific, simulated clinical scenario. 2D = 2-dimensional; 3D = 3-dimensional.

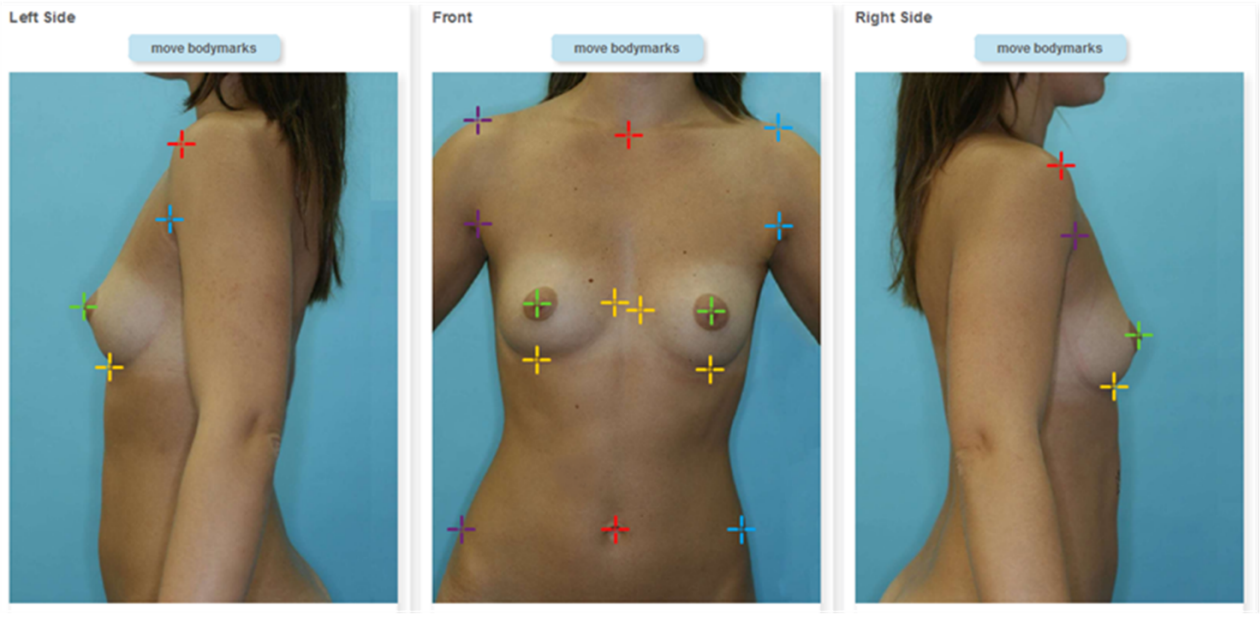
Retrieval and Analysis of Patient-Specific Information

The real-time generation of a 3D model of the patient's anatomy is based on the extraction of patient-specific information, provided by the user in the form of 2D digital pictures, taken at 3 different angles (frontal and lateral images). In addition, and in order to create a plausible model, 2 physical distance measurements of the patient's anatomy are requested, such as nipple-to-nipple and nipple-to-submammary fold. The set of 2D images and sparse measurements allow for calibration of images to the actual patient's anatomy and for reconstruction of a realistic model of the patient's anatomy.

An important aspect is to provide the user with understandable information regarding the way digital pictures need to be taken. This is indicated to the user as guidelines for taking suitable patient photos; generally, plain white hospital walls offer sufficient contrast. Conventional fluorescent lighting found in offices and hospitals is perfectly acceptable for the system to be able to reliably detect the patient's contour from the 2D pictures.

To initiate the body extraction algorithm, the user is required to define a few anatomical landmarks on each of the 3 pictures (see Figure 2). This step is guided interactively on the website and takes less than 5 minutes. From the result of the body extraction algorithm, several curvature characteristics of the breast are computed, to automatically determine the patient's breast type and to fine-tune the 3D reconstruction algorithm.

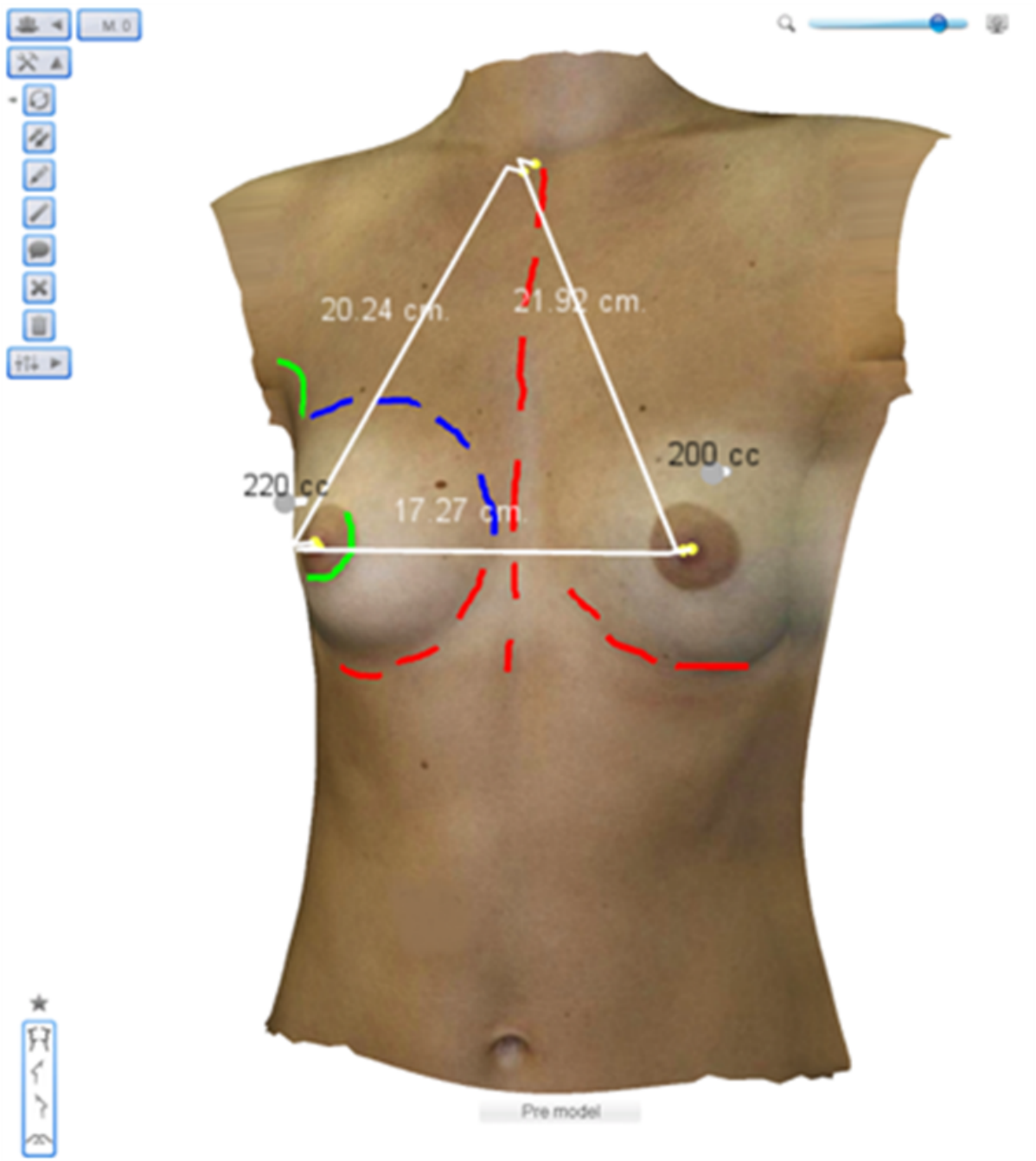
Figure 2. Three landmarked photos of a patient. Visual aids on where to place landmarks and a simple Web interface guide the user through the annotation of images. Cropped screenshot taken from the Web-based interface.



3D Reconstruction of a Patient's Anatomy and Web 3D Visualization

Once body extraction and breast type characterization is finished, a specialized image-based 3D/2D reconstruction algorithm is used to estimate the 3D shape of the patient's anatomy, from the imaging and morphometric information provided by the user. The user is presented with a 3D visual representation of the 3 views, in the form of a textured surface model (see Figure 3). The model is interactive and can be rotated by the user in the Web browser. Two technologies seem to be taking the lead in Web3D: WebGL (http://www.khronos.org/webgl/) and Unity3D (http://unity3d.com/). The first is a continuation of the O3D project from Google; despite its initial progress the engine has lost some impact due to browser incompatibilities (especially Internet Explorer) and security-related concerns raised by the community. Unity3D is a commercial product with a growing and active community, available on multiple platforms. The drawback of Unity3D is the plug-in architecture, which might hinder its dissemination to less-experienced Web users.

Figure 3. Web-based 3-dimensional (3D) annotations on a reconstructed patient model. A set of tools including 3D distances, text, and body drawing enable a personalized virtual clinical analysis. Cropped screenshot taken from the Web-based interface.



3D annotations, such as floating text, lines, and landmarks, can be added directly to the models (Figure 3). In addition, 3D measurements can be performed on the digital model of the patient, as opposed to the traditional way of painting directly

on the patient's skin. Besides the less-invasive communication approach with the patient, the clinician has the advantage of storing and accessing the 3D model on demand and without the need for the patient to be present.

Web-Enabled Biomechanical Simulations

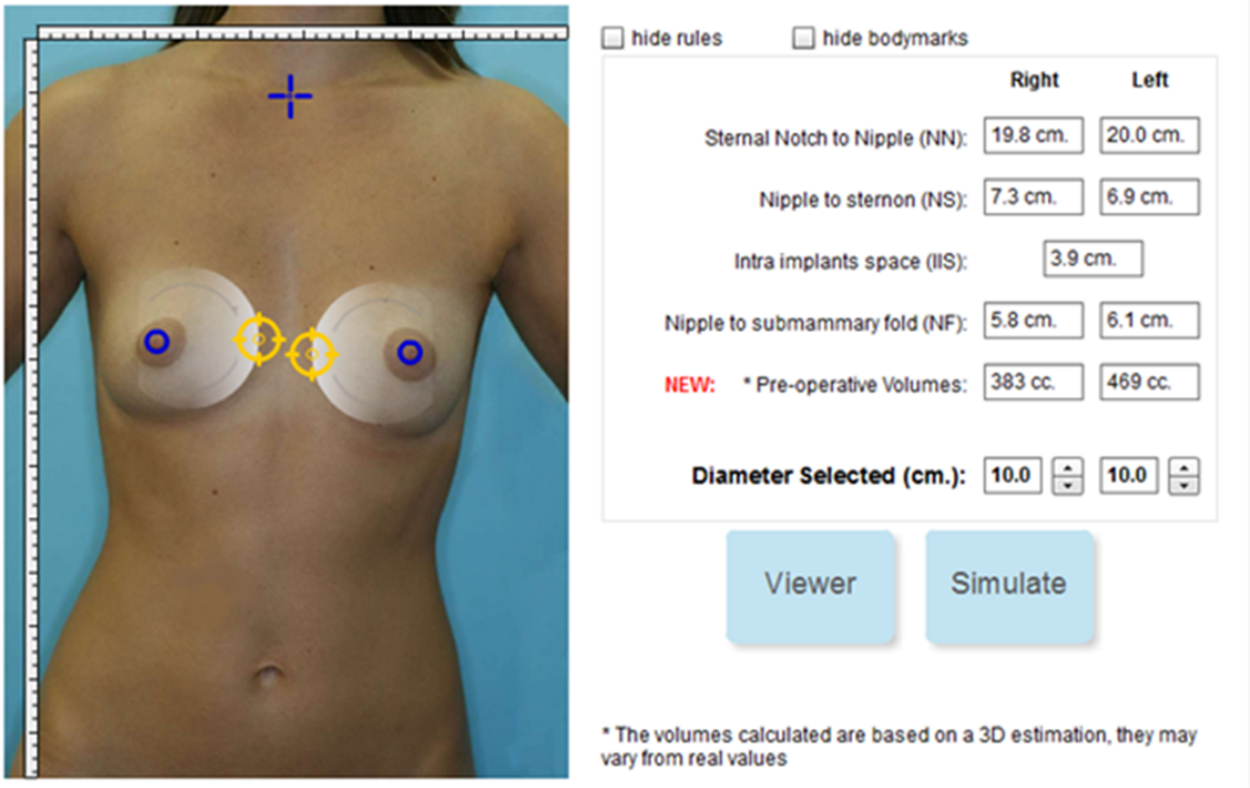
A widely accepted breast augmentation procedure consists of choosing between three different implant placement techniques, such as subglandular, submuscular, and dual plane. The implants themselves come in a plethora of different widths, heights or projections, lengths, and shapes. It is in this large array of choices that use of a physics-based implant simulator on the virtual patient is important to quickly and decisively give an idea of a final postoperative result. Furthermore, interpatient anatomical variability adds to the complex decision-making process.

Consequently, a Web-enabled simulator for breast augmentation needs to consider the current breast augmentation techniques while adapting them to the Web. The following subsections introduce the Web-based planning and biomechanical simulator.

Planning

In the planning process, implant positions and diameters are defined for use during the simulation. Positioning and sizing can be indicated directly on the photo or defined numerically (Figure 4). The surgeon can also choose the implant brand from an implant catalogue, which is a database including various brands on the market. The implant is classified according to user requirements for diameter, projection, and volume. The three methods of implantation, subglandular, submuscular, and dual plane, are specified during this step.

Figure 4. Selection of implant position and diameter. Cropped screenshot taken from the Web-based interface.



Biomechanical Simulator

To stay true to reality, many pathways have been explored in terms of viable simulation solutions such as fluidics [19], complex deformable models [20], pressure models [21], uniform mass-spring models [22], and mass tensor models [23]. First, the most important aspect of a simulation is robustness, such that a surgeon is confident it will enhance a consultation with a lifelike and consistent result. This is a challenging problem, since many parameters influence the final look, such as the shape and volume of the breasts and implants, gravity, skin tension, and the interactions between the implants and the internal tissues such as muscles, fat, glands, and skin. Second, the simulator must yield results in a reasonable time frame such that the surgeon and patient perceive the different choices in a noninterruptive manner. Third, the 3D graphical appeal of the visualization must be such that the skin, texture, and lighting conditions of the virtual reconstructed patient are smooth, so as to add to the realism of the model and emphasize the improvements to the patient's appearance.

Instead of creating a simulator that builds from prefabricated examples, our platform is based on the physical properties of human tissue using the tissue elastic model (TEM), which closely resembles the finite element method. The most important features differentiating the two methods are that in TEM

- Deformations such as torsion, volume, and angular constraints are more relaxed.
- Speed of execution is emphasized, such that simulations of thousands of iterations on large and complex aggregates of voxelized tissue (defined below) are handled in a few seconds.

- A domain-specific implementation involving a biomechanics model focuses solely on the breasts.
- Tissue elasticity is inherent in the model.

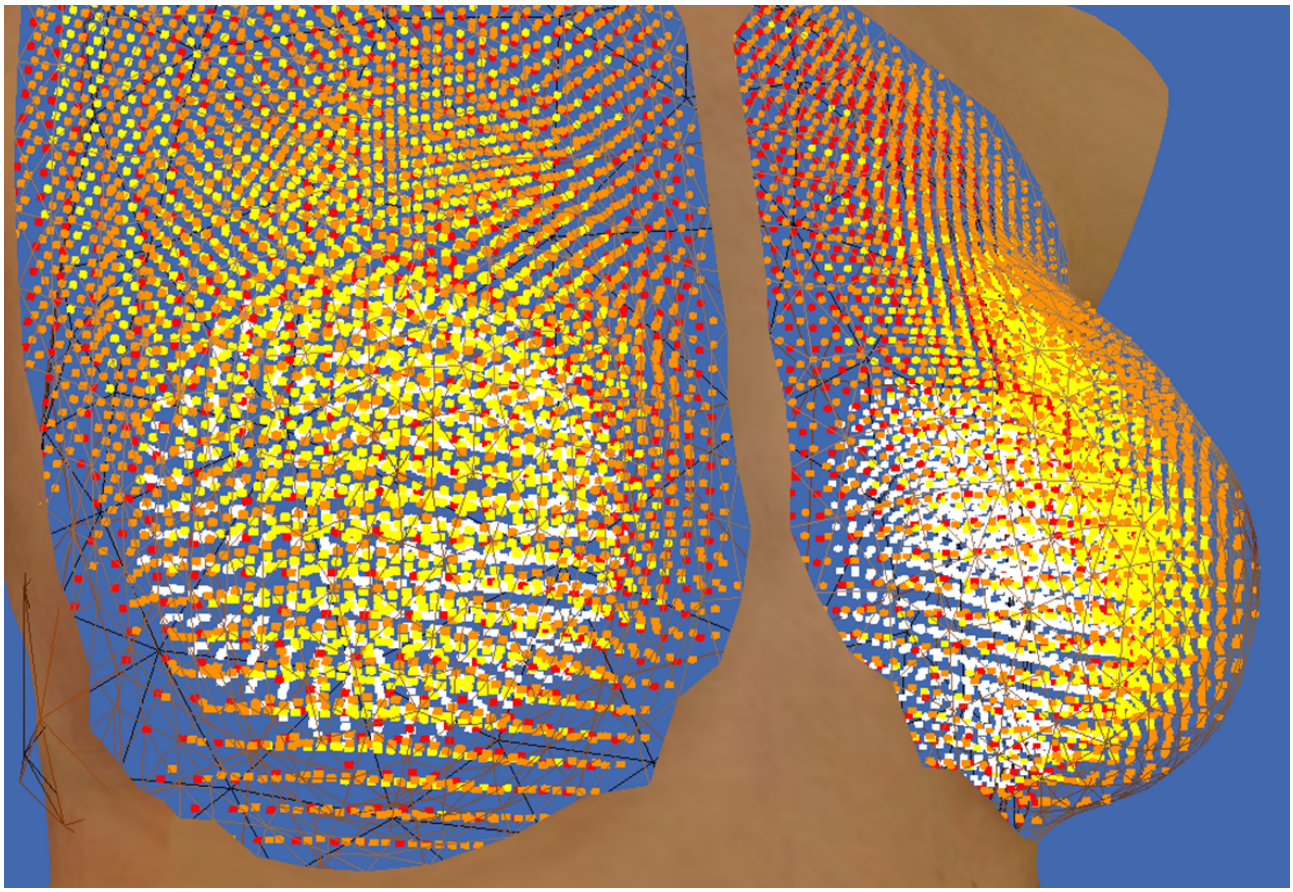
The TEM engine has an elasticity module included that takes into account the existing degree of skin elasticity and type, which is chosen by the surgeon who submits the pictures. We took into consideration 4 elasticity types: loose, moderate, tight, and very tight. Based on the selected skin elasticity the biomechanical engine modulates the global outcome of the simulation. Systems with a similar basis of gridlike structures can be found [24,25] and differ from fluid dynamics [26,27] in that neighboring relations stay the same until the final simulation step, independent of deformations.

Voxelization

The simulator starts by defining the volume constraints of the breasts and by subdividing the breast tissue into tiny 3D cubes called voxels (or volumetric pixels). This process is called voxelization and volumetrically approximates the different tissue layers of the breasts such as muscles, fat, glands, and skin. The innermost layers are the torso and bones, followed by the muscle layer of the pectorals, then fat and glands, and finally skin. This biomechanical model is medically relevant to be faithful to the surgeon's expected outcome (see Figure 5).

The simulator requires that the surgeon place the 2 implants in an image of the front part of the patient with the center part of the implant as visual guidance (see Figure 4). Once the placement is complete and the surgeon starts the simulation, the process starts by finding the space needed by the implant in the virtual patient's tissue. Depending on whether the operation is subglandular, submuscular, or dual plane, the implant is placed between the rib cage and muscle (submuscular) or between muscle and glandules, (subglandular or dual plane).

Figure 5. Voxelized breasts and implants. The fat layer is seen in yellow, the skin in orange, and the muscle layer in red. The implants are shown as white voxels or particles. This screenshot is taken from the simulation developer's point of view and is not visible in the Web-based interface.



Quality Control Assessment of the System's Precision

We retrospectively conducted a quality control assessment on available anonymized pre- and postoperative 2D digital photographs of patients undergoing breast augmentation procedures. The above-presented method was used to reconstruct 3D pictures from 2D digital pictures. A laser scanner (EScan3D [28]) capable of generating a highly accurate surface model of the patient's anatomy was used to acquire ground truth data. We compared the quality of the computed 3D reconstructions against the ground truth data used to perform both qualitative and quantitative evaluations.

For a qualitative evaluation, an overlay of the reconstruction is superimposed on the laser scan and presented to 4 plastic surgeons for direct visual comparison. The scanner is capable of submillimeter precision and can capture the breast field in 3 overlapping sweeps, each sweep lasting between 5 and 10 seconds. Slight patient motion during laser acquisition of the surfaces introduces errors in the evaluation. The patient was required to hold her breath to minimize chest motion. Eder et al [29] reviewed techniques for scanning for breast surgery and

discuss potential patient motion-related problems. Although small patient motion-induced errors are difficult to quantify accurately, they are nonetheless negligible with respect to the assumptions of the reconstruction from limited views.

Laser technology relies on several acquisitions, since the field of view is insufficient to cover the entire thorax. The patient is likely to move between acquisitions. To account for potential patient motion between scans, the patient is repositioned to stand with her back against the wall aligned with a patient-specific template with the head tilted back, and the elbow and scapula in contact with the wall. Testing of this protocol on several patients found that patient motion due to breathing was minimized, thus ensuring a good laser reconstruction. The alignment of the arms against the template also enables breast deformations to be easily matched during reconstruction of the individual 3D surface scans. These findings are also in agreement with Eder et al [29], who described virtual 3D modeling for breast surgery. Since the patient is standing with her back and elbows against the wall, she is able to relax her posture, and a better mediolateral positioning of the patient can be obtained. This is probably because the patient can rely on a spatial reference to set her posture.

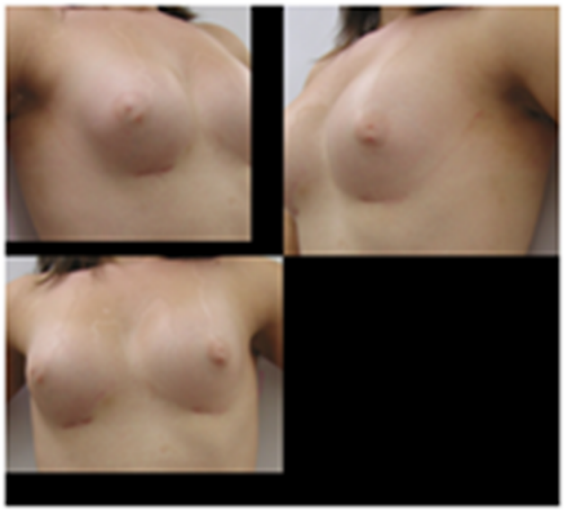
The laser scans are considered for the purposes of this study to be the criteria index, since they incorporate true depth information of the surface. To validate the algorithm, surface scans of the patient were taken at the same time as the photographs used for the 3D reconstruction. Figure 6 shows one example of the laser scanning procedure.

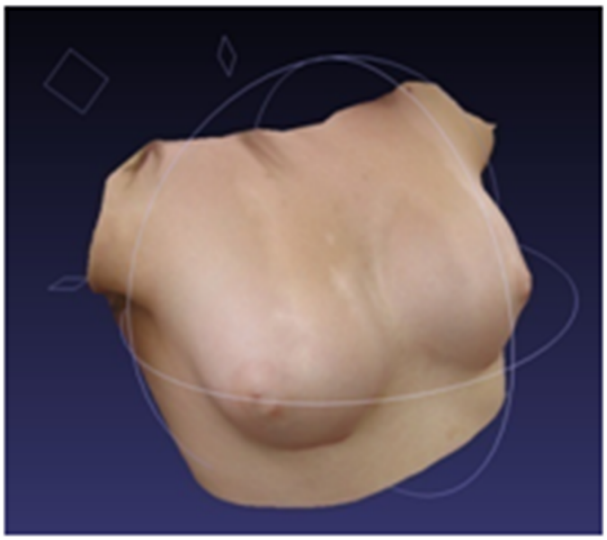
To construct an overlay of the resulting 3D reconstruction and the corresponding ground truth, the reconstructed breasts were aligned to the laser scan surface using surface-matching techniques in Amira 5.3.3 software (Visage Imaging GmbH, Berlin, Germany). The surface matching of the laser scan and the 3D reconstruction is required in order to bring meshes from different sources into the same coordinate system. During this surface-matching step, the shape of the breasts was preserved so as not to bias the results.

As a quantitative metric, the surface-to-surface distance between the overlaid reconstructed surface and the baseline scan was computed for each point of the reconstructed surface.

The quality control evaluation was based on a single ambulatory follow-up outpatient visit and had no implication in therapy planning for the patients. The current photographic 2D documentation was completed by a 3D camera scan to allow quality control of the efficiency of the Web-based 3D reconstruction system. Following anonymization of the data, the analysis was performed in a blinded manner. There are no prospective intents or implications in this retrospective photographic comparison. In total, we present 11 datasets, of which 10 are postoperative.

Figure 6. Overlapping scanner views (left) and the resulting surface scan generated by the commercial scanner software (right).





Results

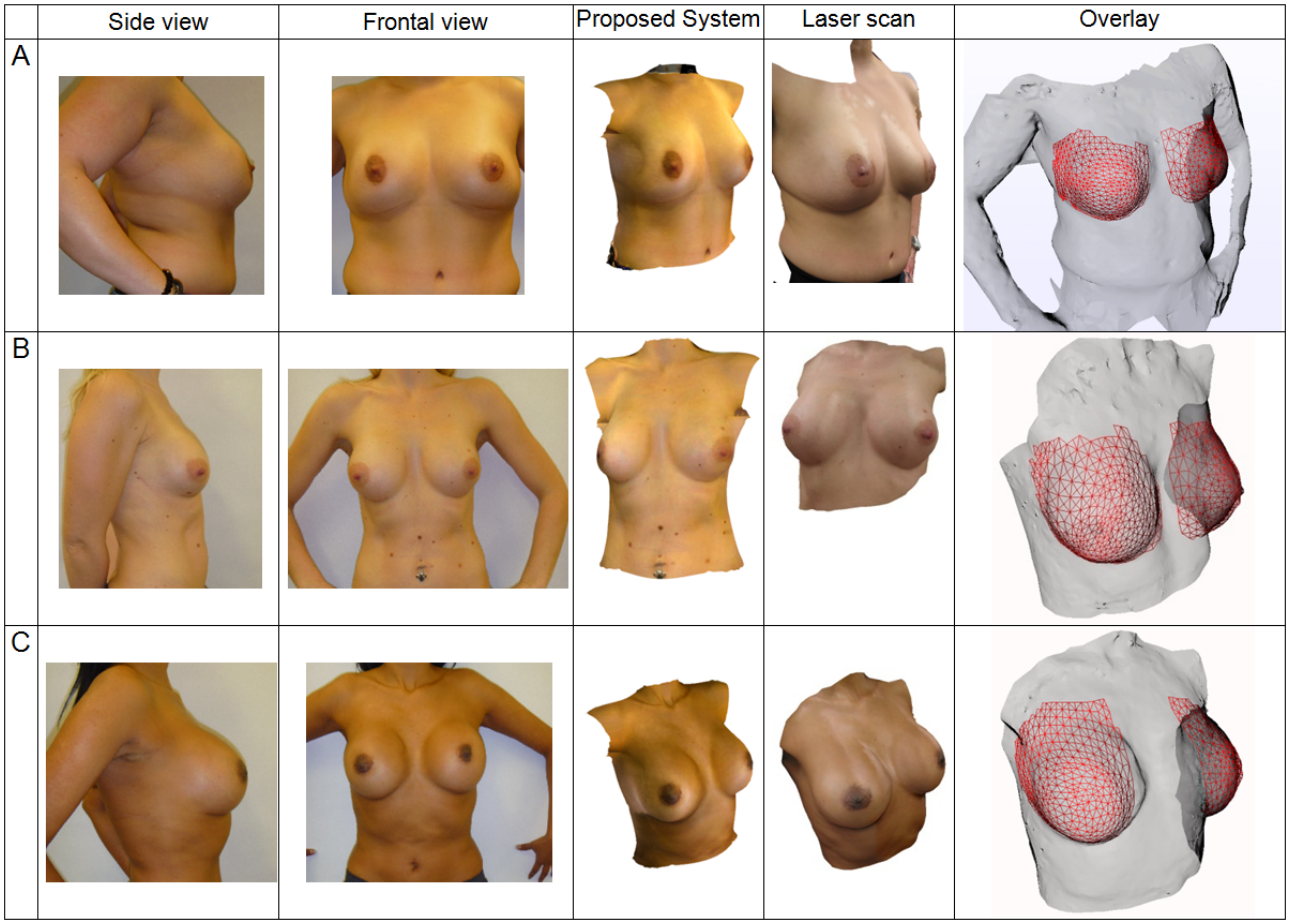
Qualitative Analysis of 3D Reconstructions of Patient Anatomies

Figure 7 is a composite figure that illustrates the corresponding patient photos, laser surface scan, and 3D reconstruction using the developed system. Due to space considerations, only the right lateral patient photograph is shown. Visually, the 3D reconstruction appears similar to the patient photos. We selected a variety of breast types (including frontal- and lateral-facing nipples) and patient thoraxes with different builds and skin tone to show the wide applicability of the method. The main restrictions in obtaining successful reconstructions are accurate patient poses according to the protocol and precisely positioned landmarks. From this composite figure, patients A and C show mismatches in texture brightness at the seams in the laser scan column. This is a limitation of the scanner software since, although it stitches the texture surfaces correctly, it does not compensate for the camera's exposure value. In comparison, the system presented in this paper is better able to handle changes in illumination than the commercial scanner's 3D reconstruction software.

To ensure that the 3D reconstruction algorithm was evaluated fairly, the breasts were cropped from the reconstructed surface

in order to directly compare against the baseline laser scan, since this is the most important part of the model to present to users. The direct comparison is illustrated by the red wire-frame mesh (reconstruction) on the gray surface (laser scan) in the last column of Figure 7.

Figure 7. Composite figure sets showing, from left to right: patient photographs, corresponding 3-dimensional (3D) surface reconstruction, laser scan ground truth, and overlaid reconstructed surfaces. The laser scan textures were acquired in the absence of flash photography; hence, their illumination appears slightly different from that in the patient photos. The laser scan surface is shown as a transparent surface in the last column, and the 3D reconstruction is displayed as a superimposed red wire frame. The preoperative simulated images are screenshots from the Web-based interface, as seen in the Proposed System column.



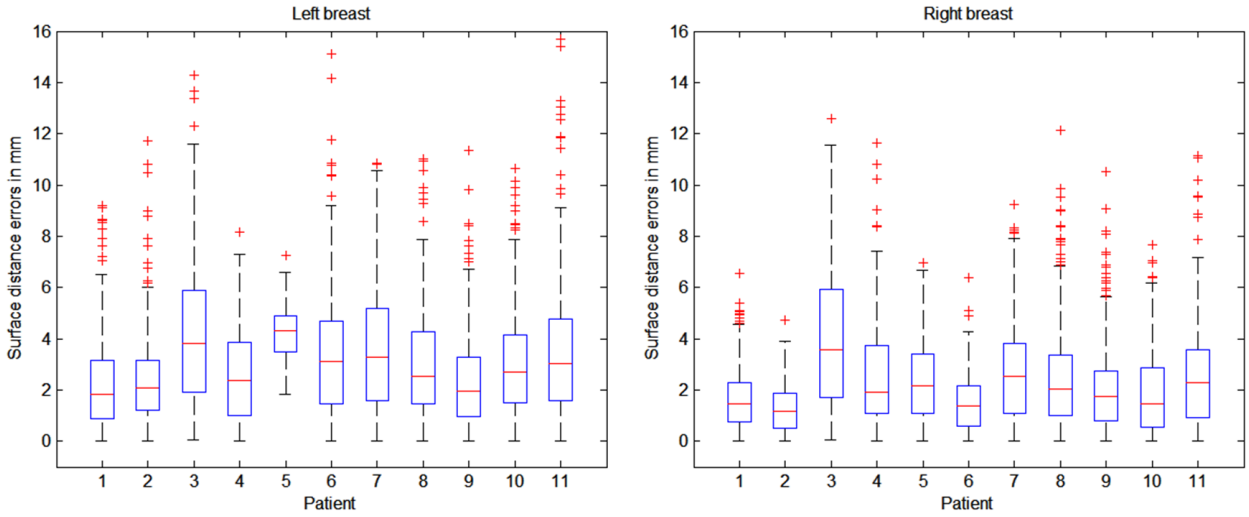
Quantitative Analysis of 3D Reconstructions of Patient Anatomies

Visually, the overlaid reconstructions appear to correlate well with the laser scan. The results of these surface distances are shown as errors on a box plot, as displayed in Figure 8. It is important to note that this surface distance analysis compares shape only and does not account for skin texture. There is a clear correlation between the box plots for the left and right breasts, with a mean reconstruction error between 2 and 4 mm for both left and right breasts. By taking into account the large 90° angle between the frontal and lateral photographs used in the reconstruction, the observed maximum surface error appears small, and is less than the motion artifacts caused by breathing excursions and small changes due to patient repositioning between scans.

A recent paper [30] highlighted the possibility that arm positioning might affect the accuracy of breast shape reconstruction using 3D laser scanning technologies. The authors suggested investigating the effect on the shape of reconstructed breasts of arms at the sides, akimbo, and akimbo with maximal force. In relation to the accuracy attained by the proposed consultation system, our experience with scanning patients in akimbo position shows only a small craniocaudal motion of the breast as compared with arms at the sides, and very small changes to the shape of the breast. In this regard, we believe that the possible artifacts introduced by akimbo positioning are negligible compared with the reported accuracy of the developed tool. A dedicated study is, however, necessary to understand and model these aspects.

It should be noted that the aim of the system is not to compete against the very accurate 3D laser scanner technologies, but to propose a Web-based patient-specific tool to aid surgeons with the consultation process and patients with their preoperative choice.

Figure 8. Box plots of left and right breast: 3-dimensional (3D) surface reconstructions compared with laser scans. The patients in Figure 7 correspond to cases 1, 7, and 8, respectively, in these box plots.



Quantitative and Qualitative Analysis of Biomechanical Simulations

To further control the quality of the proposed system as a breast implant consultation tool, simulations of postoperative surgery are predicted from the patients' preoperative photos along with the knowledge of the implants chosen for the surgery. The predictions were validated using laser scans of postoperative patients. Routine photographic documentation was taken according to common surgical practice. None of these patients were presented preoperatively with a 3D reconstruction or simulation, and no clinical therapeutic decision was based on these data.

Qualitative results are shown in Figure 9 and quantitative results are presented in Figure 10. Visually there is great variation in breast textures and sizes, which illustrates the applicability of the 3D reconstruction and simulation algorithms. The comparison of reconstructed and simulated breasts against postoperative laser scans was calculated in the same manner as in Figure 8. Just as in Figure 8, correlations between the left and right breast are apparent in Figure 10. The levels of mean errors in Figure 10 are similar to those for the reconstructions from postoperative photos in Figure 8. Since there are only 4 cases at this time for postoperative simulation, there are not enough grounds to draw strong conclusions. Nevertheless, the preliminary results indicate that the assumptions in the simulation do not dominate the overall errors. This could be because the reconstruction uses only 3 images to generate a full 3D surface and the simulation is physics based with material properties and uses knowledge of the patient-specific implants.

Figure 9. Composite figure showing pre- and postoperative photos, pre- and postoperative 3-dimensional (3D) reconstruction, and implant simulation surface renderings from the simulation visualization. The reconstructed and simulated surfaces were computed from the preoperative photos and hence show similarities in the textures. The pre- and postoperative simulated images are screenshots from the Web-based interface (middle columns).

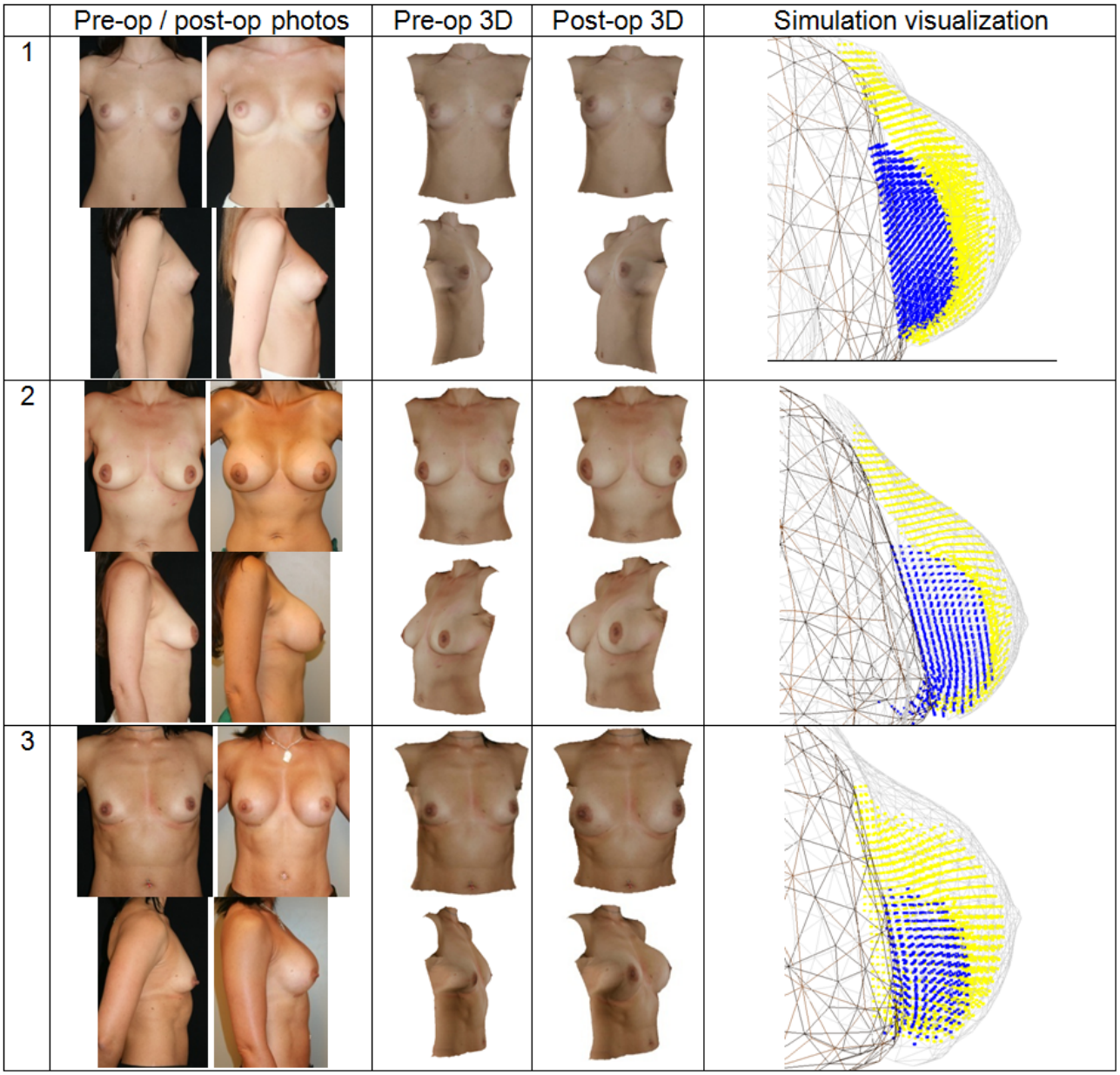
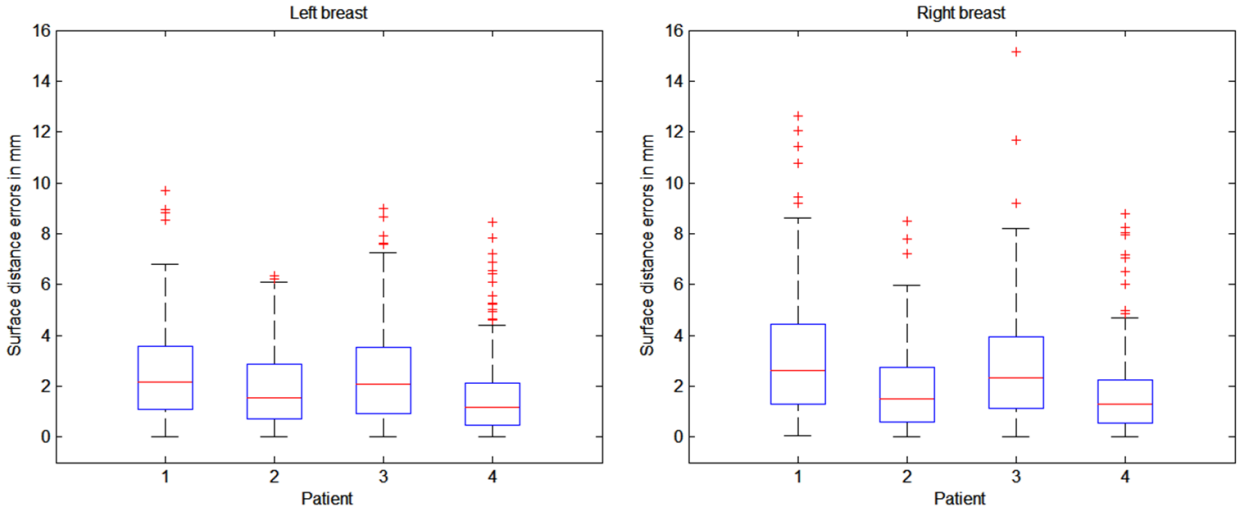


Figure 10. Postoperative simulation results predicted from preoperative images compared with postoperative laser scans.



Discussion

We present a tool developed for a 3D Web-based, patient-specific consultation in the clinical scenario of breast augmentation. The main finding of this study are that the current state of development allows for the creation of a responsive and effective Web-enabled 3D consultation tool for breast augmentation surgery based on 3D image reconstruction of 2D pictures, even with highly complex and time-consuming computation, by off-loading it to a dedicated high-performance data center.

The efficient combination of advanced technologies, based on analysis and understanding of human anatomy and physiology, will allow for the development of further Web-based reconstruction and predictive interfaces at different scales of the human body.

This consultation tool exemplifies the potential of combining advancements in the core areas of computer science and biomedical engineering along with the evolving areas of progress in Web technologies. We are confident that future developments based on a multidisciplinary approach will further pave the way toward personalized Web-enabled medicine.

Perspectives

This technology has potential for other medical applications, such as reconstructive surgery of facial malformations, aesthetic facial and anti-aging procedures, or preoperative volume evaluations in breast reduction surgery. These areas can be targeted by modeling the different anatomical and physiological processes.

For the future development of the presented system, optimization plans include improving texturing of the reconstruction from the patient's photographs, user retargeting of geometry [31], handling of complex breast shapes (congenital malformations, tubular breasts), and reconstructive scenarios (eg, breast cancer), as well as enhancing the Web2.0 interaction platforms between patients and surgeons.

Acknowledgments

We acknowledge the KTI Promotion Agency (grant number 10263.1 PFLS-LS) for their support of this study and thank Heather Gray for proofreading this paper.

Conflicts of Interest

Dr Garcia is co-owner of, and receives income from, Crisalix, S.A, which is developing products related to the research described in this paper and developed through the Swiss Agency KTI for promotion of medical technologies. The terms of this arrangement have been reviewed and approved by the University of Bern, Switzerland, in accordance with their respective conflict of interest policies.

References

1. Dwyer SJIII. A personalized view of the history of PACS in the USA. 2000 Presented at: Medical Imaging 2000: PACS Design and Evaluation: Engineering and Clinical Issues; Feb 15, 2000; San Diego, CA, USA. [doi: 10.1117/12.386388]
2. Sicurello F, Pizzi R. TME10/380: remote transmission of radiological images by means of Intranet/Internet technology. J Med Internet Res 1999;1(suppl 1):e117. [doi: 10.2196/jmir.1]
3. Menegaz G, Grewe L, Lozano A, Thiran JP. TME12/400: application oriented wavelet-based coding of volumetric medical data. J Med Internet Res 1999;1(suppl 1):e119. [doi: 10.2196/jmir.1]

4. Ferrer-Roca O, Vilarchao-Cavia J, Troyano-Luque JM, Clavijo M. Virtual sonography through the Internet: volume compression issues. J Med Internet Res 2001;3(2):E21 [FREE Full text] [doi: 10.2196/jmir.3.2.e21] [Medline: 11720963]
5. Coveney PV, Diaz V, Hunter P, Kohl P, Viceconti M. The virtual physiological human. Interface Focus 2011;1:281-285. [doi: 10.1098/rsfs.2011.0020]
6. 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] [Medline: 18725354]
7. Modiface Inc. 2009. Reshapr Plastic Surgery Visualization Tool URL: http://www.reshapr.com/ [accessed 2011-08-02] [WebCite Cache ID 60dd0HkOB]
8. KAERIA SARL. 2010. Virtual Plastic Surgery Software URL: http://www.photo-warping.com/ [accessed 2011-08-02] [WebCite Cache ID 60ddES6XC]
9. breastdoctors.com. 2011 Oct 14. Breast Augmentation URL: http://www.breastdoctors.com [accessed 2012-01-30] [WebCite Cache ID 654ahwSNm]
10. Mentor Worldwide LLC. LoveYourLook. com URL: http://www.loveyourlook.com/ [accessed 2011-08-02] [WebCite Cache ID 60ddyUzHQ]
11. Gorini A, Gaggioli A, Vigna C, Riva G. A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology. J Med Internet Res 2008;10(3):e21 [FREE Full text] [doi: 10.2196/jmir.1029] [Medline: 18678557]
12. Huang H, Rauch U, Liaw SS. Investigating learners' attitudes toward virtual reality learning environments: based on a constructivist approach. Comput Educ 2010;55(3):1171-1182. [doi: 10.1016/j.compedu.2010.05.014]
13. Nigel WJ. The impact of Web3D technologies on medical education and training. Comput Educ 2007;49(1):19-31. [doi: 10.1016/j.compedu.2005.06.003]
14. Chittaro R, Ranon R. Web3D technologies in learning, education and training: motivations, issues, opportunities. Comput Educ 2007;49(1):3-18. [doi: 10.1016/j.compedu.2005.06.002]
15. Stull A, Hegarty M, Mayer RE. Getting a handle on learning anatomy with interactive 3D graphics. J Educ Psychol 2009;101:803-816.
16. Wiecha J, Heyden R, Sternthal E, Merialdi M. Learning in a virtual world: experience with using second life for medical education. J Med Internet Res 2010;12(1):e1 [FREE Full text] [doi: 10.2196/jmir.1337] [Medline: 20097652]
17. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]
18. Dalgarno B, Lee MJW. What are the learning affordances of 3-D virtual environments? Br J Educ Technol 2010;41:10-32. [doi: 10.1111/j.1467-8535.2009.01038.x]
19. Müller M, Schirm S, Teschner M, Heidelberger B, Gross M. Interaction of fluids with deformable solids. Comput Animat Virtual Worlds 2004;15(3-4):159-171. [doi: 10.1002/cav.18]
20. Teschner M, Heidelberger B, Müller M, Gross M. A versatile and robust model for geometrically complex deformable solids. In: Proceedings. 2004 Presented at: Computer Graphics International, 2004; June 19, 2004; Crete, Greece. [doi: 10.1109/CGI.2004.1309227]
21. Matyka M, Ollila M. A pressure model for soft body simulation. In: Proceedings. 2003 Presented at: Annual SIGRAD Conference; Nov 20-21, 2003; Umeå, Sweden.
22. Rasmusson A, Mosegaard J, Sangild Sorensen T. Exploring parallel algorithms for volumetric mass-spring-damper models in CUDA. Biomed Simul 2008;5104:49-58. [doi: 10.1007/978-3-540-70521-5_6]
23. Roose L, De Maerteleire W, Mollemans W, Maes F, Suetens P. Simulation of soft-tissue deformations for breast augmentation planning. Biomed Simul 2006;4072:197-205. [doi: 10.1007/11790273_22]
24. Eitz M. Realtime Soft Tissue Simulation Employing Constraint Based Particle Systems and Hierarchical Spatial Hashing [master's thesis]. Shanghai: Jiao Tong University; 2006.
25. del Palomar AP, Calvo B, Herrero J, López J, Doblaré M. A finite element model to accurately predict real deformations of the breast. Med Eng Phys 2008 Nov;30(9):1089-1097. [doi: 10.1016/j.medengphy.2008.01.005] [Medline: 18329940]
26. Clavet S, Beaudoin P, Poulin P. Particle-based viscoelastic fluid simulation. In: Proceedings. 2005 Presented at: 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation; Jul 29-31, 2005; Los Angeles, CA, USA.
27. Solenthaler B, Pajarola R. Density contrast SPH interfaces. In: Proceedings. 2008 Presented at: 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation; Apr 14-18, 2008; Crete, Greece.
28. 3D Digital Corp. 2011. The Best in 3D Scanner Technology URL: http://www.3ddigitalcorp.com/index.shtml [accessed 2011-08-04] [WebCite Cache ID 60gjHZJuj]
29. Eder M, Papadopulos NA, Kovacs L. Re: Virtual 3-dimensional modeling as a valuable adjunct to aesthetic and reconstructive breast surgery. Am J Surg 2007 Oct;194(4):563-5; author reply 565. [doi: 10.1016/j.amjsurg.2006.11.036] [Medline: 17826081]
30. Freshwater MF. Arm position artifact in three-dimensional breast scanning technique. Plast Reconstr Surg 2011 Jul;128(1):332. [doi: 10.1097/PRS.0b013e3182173f03] [Medline: 21701366]
31. Wakita A, Hayashi T, Kanai T, Chiyokura H. Using lattice for Web-based medical applications. In: Proceedings. 2001 Presented at: Sixth International Conference on 3D Web Technology; Feb 19-22, 2001; Paderborn, Germany. [doi: 10.1145/363361.363371]

Abbreviations

2D: 2-dimensional
3D: 3-dimensional
TEM: tissue elastic model

Edited by G Eysenbach; submitted 05.08.11; peer-reviewed by M Freshwater; comments to author 19.08.11; revised version received 17.12.11; accepted 16.01.12; published 03.02.12.

Please cite as:
de Heras Ciechomski P, Constantinescu M, Garcia J, Olariu R, Dindoyal I, Le Huu S, Reyes M
Development and Implementation of a Web-Enabled 3D Consultation Tool for Breast Augmentation Surgery Based on 3D-Image Reconstruction of 2D Pictures
J Med Internet Res 2012;14(1):e21
URL: http://www.jmir.org/2012/1/e21/
doi:10.2196/jmir.1903
PMID:22306688

Original Paper

A Changing Landscape of Physician Quality Reporting: Analysis of Patients' Online Ratings of Their Physicians Over a 5-Year Period

Guodong Gordon Gao[1], PhD, MBA; Jeffrey S McCullough[2], PhD; Ritu Agarwal[1], PhD, MS, MBA; Ashish K Jha[3], MPH, MD

[1]Center for Health Information and Decision Systems, Robert H Smith School of Business, University of Maryland, College Park, MD, United States
[2]Division of Health Policy & Management, School of Public Health, University of Minnesota, Minneapolis, MN, United States
[3]Department of Health Policy and Management, Harvard School of Public Health, Harvard University, Boston, MA, United States

Corresponding Author:
Guodong Gordon Gao, PhD, MBA
Center for Health Information and Decision Systems
Robert H Smith School of Business
University of Maryland
4325 Van Munching Hall
College Park, MD, 20742
United States
Phone: 1 301 405 2218
Fax: 1 301 405 8655
Email: ggao@rhsmith.umd.edu

Abstract

Background: Americans increasingly post and consult online physician rankings, yet we know little about this new phenomenon of public physician quality reporting. Physicians worry these rankings will become an outlet for disgruntled patients.

Objective: To describe trends in patients' online ratings over time, across specialties, to identify what physician characteristics influence online ratings, and to examine how the value of ratings reflects physician quality.

Methods: We used data from RateMDs.com, which included over 386,000 national ratings from 2005 to 2010 and provided insight into the evolution of patients' online ratings. We obtained physician demographic data from the US Department of Health and Human Services' Area Resource File. Finally, we matched patients' ratings with physician-level data from the Virginia Medical Board and examined the probability of being rated and resultant rating levels.

Results: We estimate that 1 in 6 practicing US physicians received an online review by January 2010. Obstetrician/gynecologists were twice as likely to be rated ($P < .001$) as other physicians. Online reviews were generally quite positive (mean 3.93 on a scale of 1 to 5). Based on the Virginia physician population, long-time graduates were more likely to be rated, while physicians who graduated in recent years received higher average ratings ($P < .001$). Patients gave slightly higher ratings to board-certified physicians ($P = .04$), those who graduated from highly rated medical schools ($P = .002$), and those without malpractice claims ($P = .1$).

Conclusion: Online physician rating is rapidly growing in popularity and becoming commonplace with no evidence that they are dominated by disgruntled patients. There exist statistically significant correlations between the value of ratings and physician experience, board certification, education, and malpractice claims, suggesting a positive correlation between online ratings and physician quality. However, the magnitude is small. The average number of ratings per physician is still low, and most rating variation reflects evaluations of punctuality and staff. Understanding whether they truly reflect better care and how they are used will be critically important.

(J Med Internet Res 2012;14(1):e38) doi:10.2196/jmir.2003

KEYWORDS

Physician quality; online reviews; patient empowerment; quality transparency; public reporting

Introduction

There is broad consensus among policy makers and consumer groups that greater transparency in health care will improve the quality and costs of care delivered [1]. The US federal government, through the auspices of the Centers for Medicare & Medicaid Services (CMS), has led several transparency initiatives, including publicly reporting individual hospital performance on process measures and standard metrics of patient experience [2,3]. However, such efforts, when directed toward individual physicians, have been far more controversial and generally lagged behind reporting for other providers [4,5]. In the CMS's recent physician reporting initiative, individual performance measures will not be available until 2013 [6]. Policy makers and physician groups worry about having an adequate sample size to create stable estimates and not penalizing physicians who care for sicker or disadvantaged patients [5,7–9]. While national efforts at physician performance reporting have progressed slowly, a new phenomenon has begun to fill the gap.

Internet-based consumer ratings of physicians have gained interest from the private sector and are seen by many as an extension of similar user-submitted rating services, such as those focused on restaurants, hotels, books, or plumbers [10]. Advocates argue that such rating systems will provide consumers much-needed information about physician quality (at least from the consumer experience perspective) [11]. Making greater use of patient feedback is also consistent with patient empowerment, a goal set by the Affordable Care Act [12]. Critics worry that the Internet rating sites will be a forum for disgruntled patients to vent frustration over minor shortcomings, and that a small number of such ratings might tarnish physicians' reputation [13,14]. Professional societies such as the American Medical Association and even some state governments have expressed concerns about these rating programs [15,16]. Since most rating websites do not require the authentication of raters, online ratings may be subject to manipulation.

Despite these concerns, there is reason to believe that these ratings will become commonplace. Patients frequently turn to online sources for health information. While only 19% of American adults indicated that they were very likely to seek health information over the Internet in 2001 [4,8], 59% of American adults searched for health information online in 2010 [17]. Furthermore, 16% of Internet users have consulted online ratings or reviews of doctors or other providers [17,18]. This growth suggests that there is tremendous potential for these ratings to affect physician livelihood and patient behavior, but only limited studies exist [19–22]. To shed light on the growing phenomenon of online physician ratings, we used data from a major user-submitted physician review rating site to answer four questions. What proportion of US physicians have received online ratings and how has this changed over time? What types of physicians are likely to be rated? What kinds of ratings do physicians receive? And finally, are certain characteristics of physicians, such as board certification status or history of malpractice payouts, correlated with higher ratings?

Methods

Data

We developed a new dataset incorporating patients' online ratings and physician characteristics, captured from RateMDs.com. Although the number of websites that offer physician ratings has increased substantially in recent years, we focused on RateMDs.com for the following reasons. First, among the rating sites, RateMDs.com has the largest number of user-submitted reviews with narratives by a large margin, based on a recent study by Lagu et al [21] and a website ranking service [23]. Second, all the physician reviews on RateMDs.com are submitted by users voluntarily, rather than populated by surveys. There are no incentives for users to submit ratings, and the ratings are publicly available and free to use. Third, RateMDs.com started in 2004 and is one of the earliest physician review websites in the United States, while most other major competitors began rating services only after 2008. Given its early entry, RateMDs.com data enabled us to construct a 6-year period from the website's inception through January 2010 to derive insights into the historical growth trend. We retrieved all the available doctor ratings on RateMDs.com up to January 2010.

We used the Area Resource File to examine the number and distribution of US physicians, including the number of active nonfederal physicians, by specialty and geographic area. Area Resource File is a national health resource information database published by the US Department of Health and Human Services. We linked the 2009 Area Resource File data to physician ratings, by state, to better understand the growth and prevalence of patients' ratings and variations in ratings across different specialties.

To determine, in a more granular fashion, factors associated with patient rating decisions, we obtained data from the Virginia Board of Medicine. These data provide detailed information on all licensed physicians practicing in the Commonwealth of Virginia. We chose Virginia because it is a relatively large state that provides relatively detailed data on licensed physicians. We matched the ratings database to the Virginia data on the basis of name, address, and specialty. We found that the distribution of ratings across and within specialties in Virginia was very similar to that seen in the national population.

Measures

RateMDs.com's physician ratings have four domains: staff, punctuality, helpfulness, and knowledge. Patients rate physicians in each of the domains on a scale of 1 to 5, with 1 being the lowest and 5, highest. The website automatically generates an overall physician quality measure based on the average of helpfulness and knowledge ratings. While we examined all four domains, we focused on the overall quality rating as our primary outcome.

We also captured data from the website about each physician's specialty. We grouped the 97 specialties identified into five major types of physicians: primary care, medical specialists, surgeons including surgical specialists, obstetricians and gynecologists, and other specialists (such as radiologists and

pathologists). One additional rationale for separating obstetricians and gynecologists is their unique patient population (women mostly between 20 and 40 years old). We identified the location of each practice and the date of each rating. We aggregated the location data into the four broad census regions: Northeast, South, Midwest, and West.

To obtain more granular details about specific physician characteristics and their association with both the likelihood of being rated and the ratings themselves, we linked the online ratings to physicians' board certification records for the Commonwealth of Virginia. Our data from the Virginia board had 18,174 physicians, a census of Virginia's actively practicing licensed physicians as of January 2010. We matched all Virginia physicians listed on the RateMDs.com database to the Virginia licensing board data. Based on this match, we calculated that about 22% of the physician entries were duplicates due to misspelling or variation of physicians' names, which is typical for user-submitted ratings. Therefore, in all subsequent national data, we applied a correction factor (reducing the number of physicians by 22%) to calculate national statistics. We also found that patients accurately identified their physician's specialty greater than 95% of the time, suggesting that our national data are likely accurate in physician specialty designations.

We used the medical board data to determine specialty, graduation year, medical school, and malpractice claim history for each physician. We grouped the specialties into the same five major categories used for the patient rating data. After identifying the year of medical school graduation for each physician, we divided graduation years into four categories: before 1980, 1980–1989, 1990–1999, and 2000–2009. Third, we identified the medical school attended by each physician. We used a binary variable to indicate whether the medical school was ranked among the top 50 schools by *U.S.News & World Report*. Finally, we determined whether each physician had paid out any malpractice claims in the past 10 years of practice.

Analysis

We determined the total number of ratings over time for each of the five specialty categories and each geographical region. We also determined the average number of ratings per physician. Next, we examined the distribution of ratings and determined the average physician rating for each of the five specialty categories.

Among physicians practicing in Virginia, we determined the percentage who had been rated in the demographic categories of specialty, graduation year, board certification, ranking of medical school, and malpractice claim history. We used both bivariate and multivariable regression techniques to examine each of the individual characteristics and their association with the likelihood of being rated as well as the rating itself.

Results

Through January 2010, we found that more than 112,000 physicians had been rated through RateMDs.com. Compared with the national physician composition, rated physicians were more likely than the national population to be obstetrician/gynecologists and more likely to be a medical specialist (see Table 1). Male physicians were somewhat more likely to be rated (83,043/112,024, 74.13%) than would be expected based on their national composition (503,529/703,223, 71.60%), and physicians in the South were slightly more likely to be rated (39,684/112,024, 35.43%) than their national composition (233,332/703,223, 33.18%).

Table 1. Comparison of the physicians rated in RateMDs.com versus the national physician population.

	Rated physicians (n = 112,024)		National physicians[a] (n = 703,223)	
Specialty				
Primary care	45,552	40.66%	280,273	39.86%
Medical specialties	17,754	15.85%	72,073	10.25%
Surgeon/surgical specialties	22,657	20.23%	113,011	16.07%
Obstetrics/gynecology	12,978	11.59%	40,013	5.69%
Other specialties	13,083	11.68%	197,853	28.14%
Gender				
Male	83,043	74.13%	503,529	71.60%
Female	28,981	25.87%	199,694	28.40%
Region				
Northeast	25,663	22.91%	168,600	23.98%
South	39,684	35.42%	233,332	33.18%
West	23,742	21.19%	153,552	21.84%
Midwest	22,935	20.47%	147,739	21.01%

[a] Based on 2007 Area Resource File patient-care physician data.

Physician Rating Trends

From the inception of RateMDs.com in 2004, there has been strong enthusiasm from voluntary users. By January 31, 2010, after excluding non-US and nonphysician practitioners, there were a total of 368,559 physician ratings (Figure 1), a 100-fold increase in the preceding 5 years. The number of individual physicians who had at least one online rating grew commensurately from 2475 physicians in January 2005 to 112,024 by January 2010, covering about 16% of all practicing US physicians.

The likelihood of being rated varied widely across specialties as of January 2010: while 32.43% (12,978/40,013) of all obstetrician/gynecologists had been rated, approximately 24.63% (17,754/72,073) of medical specialists, 20% (22,657/113,011) of surgeons, and 16.25% (45,552/280,273) of primary care physicians had received a rating. Only 6.61% (13,083/197,853) of physicians classified as other specialists (such as radiologists, pathologists, and anesthesiologists) had been rated (Figure 2). This pattern is consistent across the regions.

The average number of ratings per physician in January 2010 was 3.2, and nearly half of the physicians had only a single rating. However, the number of physicians with five or more ratings rose rapidly from less than 1% in 2005 to 12.50% (14,003/112,024) in 2010.

Figure 1. Cumulative number of ratings and rated physicians (based on data from RateMDs.com).

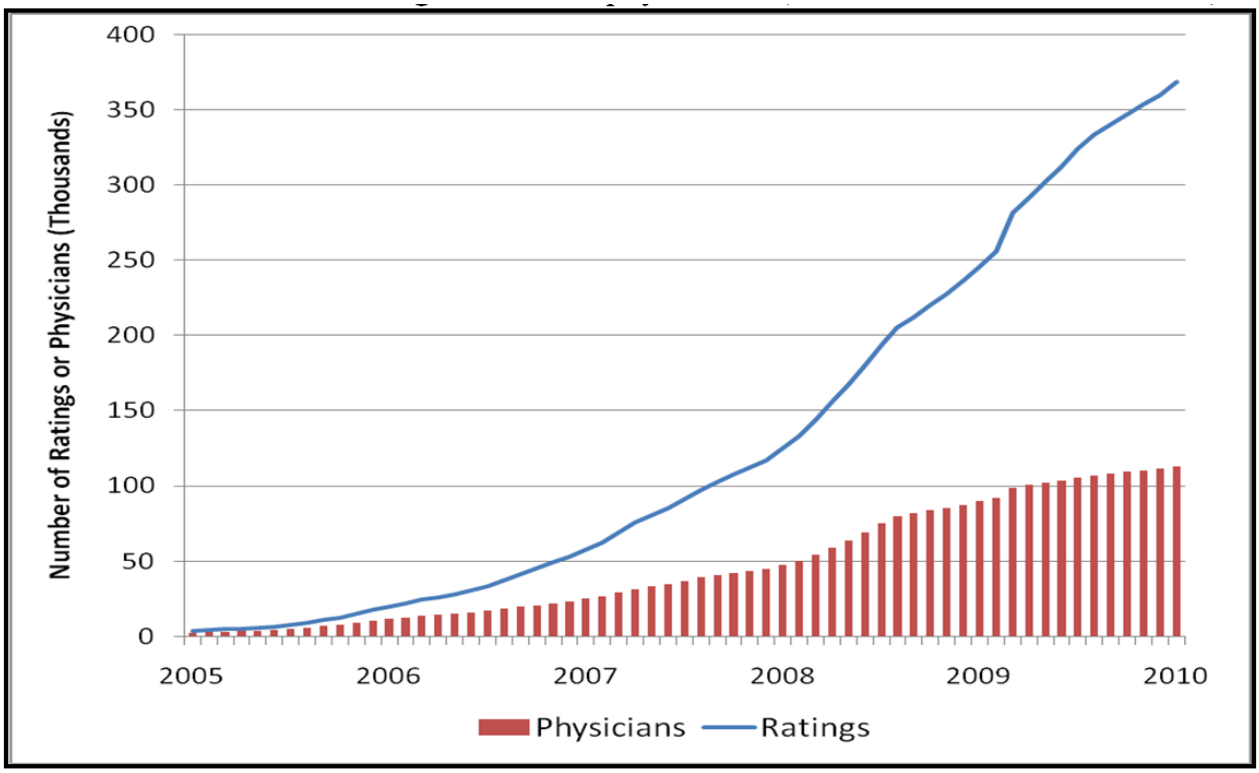
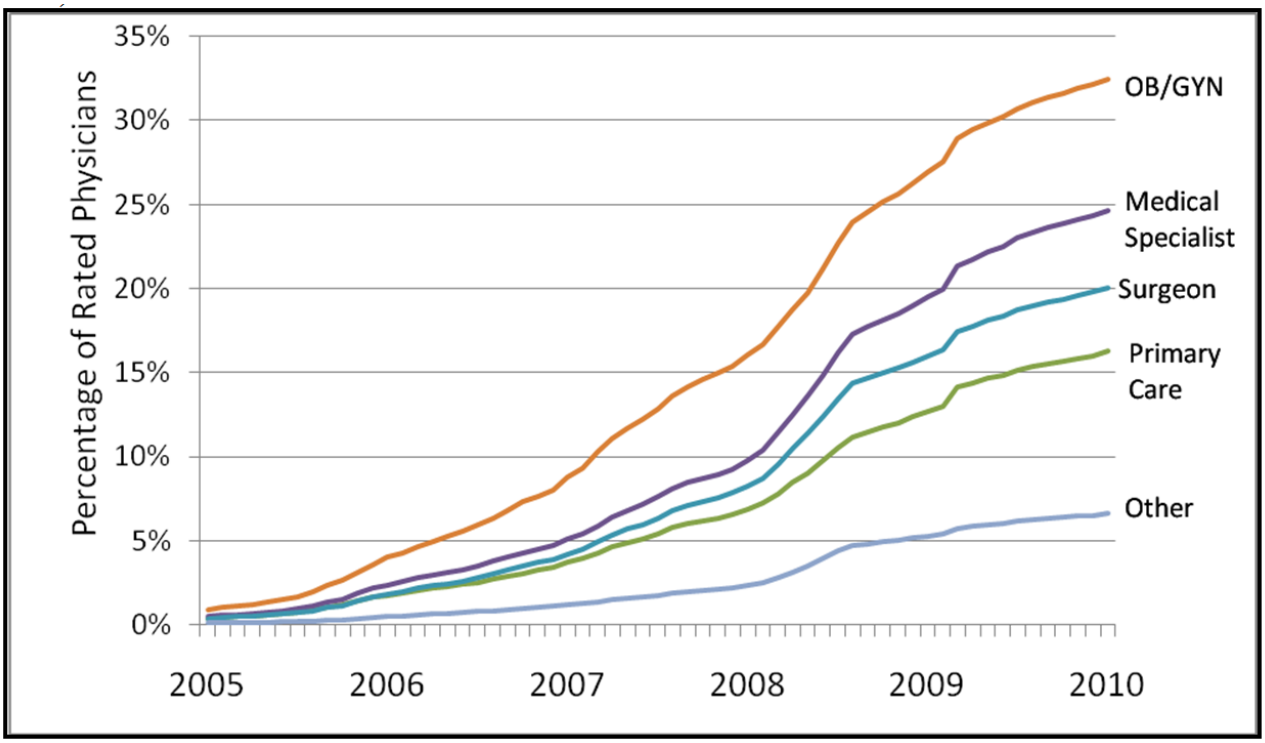


Figure 2. Percentage of US physicians rated in each specialty (based on data from RateMDs.com and the US Department of Health and Human Services' Area Resource File). OB/GYN = obstetrician/gynecologists.

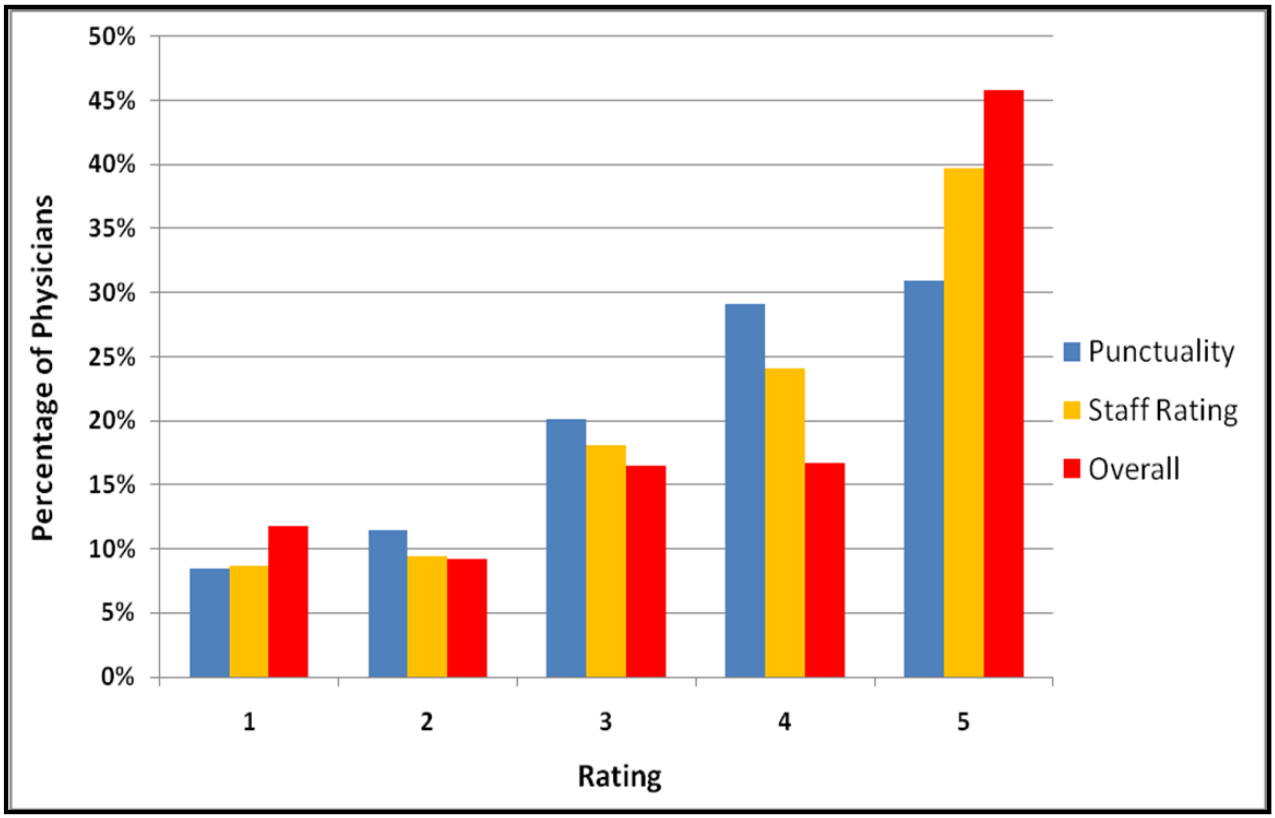


Distribution of Ratings

The average quality rating, which is based on the physician's helpfulness and knowledge, was high (3.93 out of 5), with 45.80% (51,307/112,024) of physicians receiving a 5 out of 5, while only 11.76% (13,174/112,024) of ratings were below 2 (Figure 3). Ratings on the two other dimensions (quality of staff and punctuality of physician), while still generally positive, were somewhat lower: only 30.88% (34,593/112,024) of ratings for punctuality were a 5 out of 5, while 39.68% (44,451/112,024) of the ratings for staff helpfulness were a 5 out of 5 (Figure 3). The correlation between physician quality rating and staff rating was .73 ($P < .001$), and the correlation between quality and punctuality was .68 ($P < .001$), consistent with a previous finding [22].

Across specialties, we found that the mean quality ratings were similar for physicians in primary care (4.02), medical specialties (3.96), surgeon and surgical specialties (3.89), and obstetrician/gynecologists physicians (4.01). Physicians listed within the group of other specialties had lower ratings (3.59), a difference that was statistically significant (data not shown). Finally, male physicians had, on average, a somewhat higher rating than female physicians (3.95 vs 3.89, $P < .001$).

Figure 3. Distribution of quality ratings across physicians.



Virginia Physicians' Likelihood of Being Rated

Of the 18,174 physicians in Virginia, 3164, or 1 out of every 6, had received at least one rating, which was similar to the national rate. These Virginia physicians had a total of 10,534 ratings by January 31, 2010, with an average rating of 3.3 per physician, also consistent with the national data. Much like in the national data, obstetrician/gynecologists were far more likely than others to be rated (see Table 2). Younger physicians (those graduating from medical school after 2000) were much less likely to be rated. We found that physicians who were board certified and those who had at least one paid malpractice claim were more likely to be rated, although the difference based on malpractice claim did not reach statistical significance ($P = .12$). Graduates of more highly ranked medical schools were rated with nearly the same frequency as graduates of lower-ranked medical schools. When we used multivariable models to examine the association of these factors with the likelihood of being rated, we found nearly identical results (see Multimedia Appendix 1, Table 1).

Table 2. Virginia physicians by quality rating status.

	Virginia physicians	Unadjusted rate of being rated (%)	*P* value[a]
Number of physicians	18,174	17	
Specialty			<.001
Primary care	6540	19	
Medical specialties	2806	20	
Surgeon/surgical specialties	2751	22	
Obstetrics/gynecology	1145	37	
Other specialties	4932	7	
Graduation year			<.001
Before 1980	5142	17	
1980–1989	5276	20	
1990–1999	5184	20	
2000–2009	2572	9	
Board certification			<.001
Board certified	15,057	19	
Not board certified	3117	10	
Medical school ranking			.42
Ranked top 50[b]	4962	18	
Ranked below top 50	13,212	17	
Malpractice claims			.12
No malpractice claims	16,886	17	
At least one malpractice claim	1288	23	

[a] Based on joint χ^2 test for subgroup differences with other controls.

[b] Based on 2008 *U.S.News & World Report* ranking.

We found modest effects of specialty on the quality rating that physicians received. Consistent with our national data, Virginia physicians classified as practicing other specialties had moderately lower ratings (3.63) than other physician specialty categories (Table 3). The differences between the other specialists, such as primary care physicians (4.04), medical specialists (3.95), surgeons (3.90), and obstetrician/gynecologists (4.04) were small and not significant.

Younger physicians—those who had graduated from medical school after 2000—had significantly higher ratings than older physicians. While there were small differences across the different age cohorts (3.85 for physicians graduating before 1980, 3.95 for those graduating in the 1980s, and 3.99 for those graduating in the 1990s), the youngest cohort had an average rating of 4.22 ($P < .001$ for differences across groups). Board-certified physicians had somewhat higher ratings than physicians who were not board certified (3.96 vs 3.86, $P = .04$). Similarly, physicians graduating from a top-50 medical school had somewhat higher ratings than other physicians (4.08 vs 3.91, $P = .002$). Physicians with no history of paying malpractice claims were rated somewhat higher than physicians who had at least one malpractice claim, although this difference did not reach statistical significance ($P = .1$). Once again, we found very similar results in our multivariable models (see Multimedia Appendix 1, Table 2).

Table 3. Virginia physicians by the value of quality rating.

	Unadjusted average rating	*P* value[a]
All rated physicians	3.95	
Specialty		<.001
Primary care	4.04	
Medical specialties	3.95	
Surgeon/surgical specialties	3.90	
Obstetrics/gynecology	4.04	
Other specialties	3.63	
Graduation year		<.001
Before 1980	3.85	
1980–1989	3.95	
1990–1999	3.99	
2000–2009	4.22	
Board certification		.04
Board certified	3.96	
Not board certified	3.86	
Medical school ranking in primary care		.002
Ranked top 50[b]	4.08	
Ranked below top 50	3.91	
Malpractice claims		.1
No malpractice claims	3.97	
At least one malpractice claim	3.82	

[a] Based on joint *F* test for subgroup differences with other controls.

[b] Based on 2008 *U.S.News & World Report* ranking.

Discussion

We examined physician ratings on a major user-submitted physician review website in the United States. We found dramatic growth in the number of physicians being rated (1 in 6 practicing doctors in the nation in January 2010). This trend was widespread across the nation, and its penetration differed based on physician specialty. Not surprisingly, physicians with less direct patient contact (such as pathologists or radiologists) were infrequently rated. On the other hand, obstetrician/gynecologists were far more likely to be rated than others with nearly 1 in 3 such physicians now having an online rating at RateMDs.com. This could be attributed to the younger and female patient population, who are more likely to be active Internet users. However, the difference between specialties in the propensity to be rated deserves future study. We found that ratings were generally quite positive; however, patients were more critical in their ratings of staff and punctuality.

Although some physicians are concerned that online ratings will become a channel for disgruntled patients to vent their complaints [13,14], our findings suggest that this is not the most common reason patients use the rating system. In fact, given that nearly half the physicians received a perfect 5 out of 5, online ratings appear to be driven by patients who are delighted with their physicians. Conversely, only 12% of ratings were below 2. These findings should allay concerns that online rating sites disproportionately attract dissatisfied patients; however, the prevalence of high ratings may reflect ratings selection regarding which physicians are rated.

One major concern with the physician ratings is the small sample sizes. The average number of ratings per physician is only three, and approximately half of all physicians were rated by only one patient. This low density of ratings makes physicians' average ratings vulnerable to large swings from the input of a single patient or manipulation from providers.

By linking the ratings with Virginia Board of Medicine licensing data, we found that board-certified physicians as well as those who attended higher-ranked schools had better ratings, although the differences were small. There are several potential explanations. First, it is possible that patients were aware of their physicians' school rankings or their board certification status and were, therefore, biased toward being more favorable. Alternatively, board-certified physicians from highly ranked schools might care for a patient population more positively predisposed to ranking their physicians highly. While these are all plausible, it is also possible that patients are adept at identifying high-quality physicians, and their rankings reflect the generally better care that these physicians might provide.

Certainly, prior studies have found that patients rate their experiences more highly in hospitals that have higher-quality performance [3,24], and our findings might reflect a similar phenomenon in the ambulatory setting. The finding that physicians with a history of paid malpractice claims had slightly lower ratings further supports this explanation. By linking individual physician characteristics to their online ratings, this study provides the first statistical evidence on how online ratings by patients are associated with physician quality. Further work is needed to discern whether these ratings are correlated with actual clinical outcome.

The notion that patients will use the Internet for medical information should come as no surprise. Some studies have found that 61% of US adults have looked online for health information, and among them 24% have consulted rankings or reviews online of doctors or other providers [17]. Given that patients place substantial trust in online health information (with some studies suggesting, surprisingly, that patients may place more trust in Internet-based health information than opinions of friends or coworkers [18]), the use of these kinds of data for choosing providers will likely become commonplace.

Indeed, the potential use of these data by consumers is a concern that has led some organizations to speak out against these rating programs, worried that negative ratings might harm physician practice volume and could affect physician livelihood [25]. Others, such as the National Health Service in the United Kingdom, have taken a different tack: starting October 2009, they are encouraging patients to rate their general practitioners through a National Health Service-run website [26]. Lagu and Lindenauer recently called on the CMS to put the public back in public reporting by allowing consumers to report their experiences on Hospital Compare and other government-run websites that feature provider quality performance [11].

A limitation of our work is that we aggregated data from a single website. Although it is ranked highly among physician review websites [23], in recent years, several other websites including healthgrades.com, vitals.com, yelp, and Angie's List have been offering physician ratings and are growing in popularity. Lagu et al identified 33 websites that provide physician ratings [21]. However, most of the major competitors did not start physician ratings until very recently, and some report patient survey results rather than patient-initiated reviews. Nevertheless, the values we generated for the number of physicians who have an online reputation almost certainly understate the phenomenon. Unfortunately, aggregating data across multiple websites is impractical given the potential number of practicing US physicians, the growing number of websites that offer the physician rating service, and incompatibility in rating methods and review collection approaches [20,22]. Our estimates therefore provide a lower bound of the magnitude of this important phenomenon. Future studies could examine patient choice of different rating websites.

Another limitation of our work is that our data are based on a limited number of ratings for physicians, although in aggregate, they reflect the current state of the ratings program. Further, given that the phenomenon of physician ratings is relatively new, we suspect that as more patients rate their physicians, the associations we found may change over time. Additionally, since the website could not verify the identities of the authors of reviews, it is possible that some ratings were subject to manipulation. Although the website has taken certain actions, including disallowing multiple ratings for a physician from the same computer, and removing self-promoting reviews once detected, the possibility of manipulation cannot be completely eliminated. Finally, our physician characteristics data came from a single state, and although the state appeared to be representative in the ways we could measure (ie, percentages of physicians rated, the rating scores themselves, etc), it is unclear whether the associations found using Virginia data generalize to the rest of the nation. We also lack finer measures of physician quality to be associated with the online ratings.

In conclusion, this study makes unique contributions to our understanding of online doctor ratings by examining its national growth trend, and by identifying influential factors such as specialty, board certification, education, and malpractice claims. We found that ratings were generally positive, and certain types of clinicians were both more likely to be rated and to be rated more highly. Whether the medical community and policy makers are supportive of this phenomenon or not, our findings suggest that user-generated reviews of providers are here to stay and likely to grow. We also found a weak correlation between the online ratings and physician quality. This ought to create greater impetus for policy makers to provide a context for user-generated data by speeding up current efforts to report physician quality scores online. The explosion of this information suggests that consumers are not only generating these data but likely also using it. Given the potential impact of this phenomenon, there is a great need to examine the information value and potential biases inherent in Web-based ratings, how they are being used, the impact they are having in physician choice (if any), and how to help patients make the best use of the online ratings to complement other physician performance information [27].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Statistical appendix.

[PDF File (Adobe PDF File), 87KB - jmir_v14i1e38_app1.pdf]

References

1. Agency for Healthcare Research and Quality. 2011 Aug 17. Public Reporting as a Quality Improvement Strategy: A Systematic Review of the Multiple Pathways Public Reporting May Influence Quality of Health Care URL: http://www.effectivehealthcare.ahrq.gov/ehc/products/343/763/CQG-Public-Reporting_Protocol_20110817.pdf [accessed 2012-02-20] [WebCite Cache ID 65bBjIRE9]
2. Jha AK, Li Z, Orav EJ, Epstein AM. Care in U.S. hospitals--the Hospital Quality Alliance program. N Engl J Med 2005 Jul 21;353(3):265-274 [FREE Full text] [doi: 10.1056/NEJMsa051249] [Medline: 16034012]
3. Jha AK, Orav EJ, Zheng J, Epstein AM. Patients' perception of hospital care in the United States. N Engl J Med 2008 Oct 30;359(18):1921-1931 [FREE Full text] [doi: 10.1056/NEJMsa0804116] [Medline: 18971493]
4. Harris KM, Buntin MB. Robert Wood Johnson Foundation. 2008 May. Choosing a Health Care Provider: The Role of Quality Information URL: http://www.rwjf.org/files/research/051508.policysynthesis.qualityinfo.rpt.pdf [accessed 2012-02-20] [WebCite Cache ID 65bBkAtQi]
5. Jha AK, Epstein AM. The predictive accuracy of the New York State coronary artery bypass surgery report-card system. Health Aff (Millwood) 2006;25(3):844-855. [doi: 10.1377/hlthaff.25.3.844] [Medline: 16684751]
6. Centers for Medicare and Medicaid Services (CMS). 2011 Sep 20. Physician Compare Web Site Overview: Information for Professionals URL: http://www.cms.gov/Physician-Compare-Initiative/ [accessed 2012-02-20] [WebCite Cache ID 65bBnw4pk]
7. Marcin JP, Li Z, Kravitz RL, Dai JJ, Rocke DM, Romano PS. The CABG surgery volume-outcome relationship: temporal trends and selection effects in California, 1998-2004. Health Serv Res 2008 Feb;43(1 Pt 1):174-192. [doi: 10.1111/j.1475-6773.2007.00740.x] [Medline: 18211524]
8. Shahian DM, Normand SL, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE, et al. Cardiac surgery report cards: comprehensive review and statistical critique. Ann Thorac Surg 2001 Dec;72(6):2155-2168. [Medline: 11789828]
9. Christianson JB, Volmar KM, Alexander J, Scanlon DP. A report card on provider report cards: current status of the health care transparency movement. J Gen Intern Med 2010 Nov;25(11):1235-1241. [doi: 10.1007/s11606-010-1438-2] [Medline: 20625849]
10. Fuhrmans V. Wall Street Journal Health Blog. 2007 Oct 22. Zagat Gets Into Doctor Ratings URL: http://blogs.wsj.com/health/2007/10/22/zagat-gets-into-doctor-ratings-business/ [accessed 2012-02-20] [WebCite Cache ID 65bBraYmt]
11. Lagu T, Lindenauer PK. Putting the public back in public reporting of health care quality. JAMA 2010 Oct 20;304(15):1711-1712. [doi: 10.1001/jama.2010.1499] [Medline: 20959582]
12. HHS Press Office. US Department of Health and Human Services. 2011 Apr 29. Administration Implements Affordable Care Act Provision to Improve Care, Lower Costs URL: http://www.hhs.gov/news/press/2011pres/04/20110429a.html [accessed 2012-02-15] [WebCite Cache ID 65TgaQwk7]
13. McCartney M. Will doctor rating sites improve the quality of care? No. BMJ 2009;338:b1033. [Medline: 19293224]
14. Jain S. Googling ourselves--what physicians can learn from online rating sites. N Engl J Med 2010 Jan 7;362(1):6-7 [FREE Full text] [doi: 10.1056/NEJMp0903473] [Medline: 20054044]
15. O'Connor D. American Medical News. Rated Negatively Online? What's a Physician to Do? 2011 Aug 8 URL: http://www.ama-assn.org/amednews/2011/08/08/prca0808.htm [accessed 2011-11-17] [WebCite Cache ID 63HL77Ikt]
16. Martin R. To doctors' dismay, web ratings grow. St Petersburg Times May 17 2009.
17. Fox S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2011 May 12. URL: http://pewinternet.org/~/media//Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2012-02-20] [WebCite Cache ID 65bBycrzU]
18. Elkin N. How America Searches: Health and Wellness.: iCrossing; 2008. URL: http://www.icrossing.com/sites/default/files/how-america-searches-health-and-wellness.pdf [accessed 2012-02-20] [WebCite Cache ID 65bC192iK]
19. Black EW, Thompson LA, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. Inform Prim Care 2009;17(4):249-253. [Medline: 20359403]
20. Reimann S, Strech D. The representation of patient experience and satisfaction in physician rating sites. A criteria-based analysis of English- and German-language sites. BMC Health Serv Res 2010;10:332 [FREE Full text] [doi: 10.1186/1472-6963-10-332] [Medline: 21138579]
21. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946. [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
22. Kadry B, Chu LF, Kadry B, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. J Med Internet Res 2011;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
23. Xmarks, Inc. 2011. Site: www.ratemds.com URL: http://www.xmarks.com/site/www.ratemds.com/ [accessed 2011-11-17] [WebCite Cache ID 63HLxuzKR]
24. Kaiser Family Foundation, Agency for Healthcare Research and Quality, and Harvard School of Public Health. 2004 Nov. National Survey on Consumers' Experiences With Patient Safety and Quality Information URL: http://www.kff.org/kaiserpolls/pomr111704pkg.cfm [accessed 2012-02-20] [WebCite Cache ID 65bC7UKi3]

25. ElBoghdady D. The Washington Post. 2012 Jan 28. Some Doctors Try to Squelch Online Reviews URL: http://www.washingtonpost.com/business/economy/some-doctors-try-to-squelch-online-reviews/2011/11/29/gIQA2KQhYQ_story.html [accessed 2012-02-20] [WebCite Cache ID 65bC9nTfN]
26. Smith R. The Daily Telegraph. 2009 Oct 14. Patients Can Post Online Comments About GP Practice URL: http://www.telegraph.co.uk/health/healthnews/6316125/Patients-can-post-online-comments-about-GP-practice.html [accessed 2012-02-20] [WebCite Cache ID 65bCBF2TS]
27. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]

Abbreviations

CMS: Centers for Medicare & Medicaid Services

Edited by G Eysenbach; submitted 18.11.11; peer-reviewed by W Duan, G Paré, B Gu, D Strech, J Hao; comments to author 06.12.11; revised version received 15.02.12; accepted 15.02.12; published 24.02.12.

Please cite as:
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
URL: http://www.jmir.org/2012/1/e38/
doi:10.2196/jmir.2003
PMID:22366336

Original Paper

Online Availability and Safety of Drugs in Shortage: A Descriptive Study of Internet Vendor Characteristics

Bryan A Liang[1,2], MD, JD, PhD; Tim K Mackey[1,3], MAS

[1]Institute of Health Law Studies, California Western School of Law, San Diego, CA, United States

[2]San Diego Center for Patient Safety, Department of Anesthesiology, University of California, San Diego - School of Medicine, San Diego, CA, United States

[3]Joint Doctoral Program in Global Health, University of California, San Diego - San Diego State University, San Diego, CA, United States

Corresponding Author:
Bryan A Liang, MD, JD, PhD
Institute of Health Law Studies
California Western School of Law
350 Cedar Street
San Diego, CA, 92101
United States
Phone: 1 619 515 1568
Fax: 1 619 515 1599
Email: baliang@alum.mit.edu

Abstract

Background: Unprecedented drug shortages announced by the US Food and Drug Administration (FDA) have severely affected therapeutic access, patient safety, and public health. With continued shortages, patients may seek drugs online.

Objective: To assess the prevalence of online marketing for current FDA shortage drugs and potential patient safety risks.

Methods: We performed a descriptive study of the prevalence of online marketing for shortage drugs—that is, offers for sale of each drug, including characteristics of online drug sellers and intermediary sites marketing these drugs.

Results: Of the 72 FDA shortage-listed drugs, 68 (94%) were offered for sale online. We found 291 offers for these drugs, the vast majority (n = 207, 71.1%) by online drug sellers selling direct to consumers. Intermediary sites included data aggregators (n = 22, 8%), forum links (n = 23, 8%), and personal page data links (n = 34, 12%), as well as Flickr social media links (n = 5, 2%), all advertising drugs without a prescription. Of the 91 online drug sellers identified, 31 (34%) had more than 1 shortage drug offered for sale, representing most (n = 148, 71%) of all online drug seller sales offers. The majority of these online drug sellers (n = 21, 68%) were on the National Association of Boards of Pharmacy (NABP) Not Recommended Sites list. Finally, for shortage drugs with an online drug seller (n = 58, 85%), 53 (91%) had at least one site on the Not Recommended list and 21 (36%) had only sites on the Not Recommended list.

Conclusions: FDA shortage drugs are widely marketed over the Internet. Suspect online drug sellers and intermediaries dominate these sales offers. As a critical risk management issue, patients, providers, and policymakers should be extremely cautious in procuring shortage drugs through Internet sourcing.

(J Med Internet Res 2012;14(1):e27) doi:10.2196/jmir.1999

KEYWORDS

Illicit online pharmacies; FDA drug shortage; gray market; drug supply; social media; direct-to-consumer advertising; Internet pharmacies; health policy

Introduction

The US Food and Drug Administration (FDA), the American Hospital Association, and the American Society of Health-System Pharmacists have announced a growing crisis of drug shortages in critical areas such as cancer therapy, urgent care, infectious disease, hypertension, orphan disease, pediatric, and other key patient and disease treatment categories [1-6]. According to the Premier healthcare alliance, over the last 6 months of 2010, 89% of providers experienced shortages that may have caused a medication safety issue or error in patient care, 80% experienced shortages that resulted in a delay or cancellation of a patient care intervention, and 98% experienced

shortages that resulted in increased costs [7]. A survey by the American Hospital Association found similarly dramatic results, including 99.5% of surveyed hospitals reporting a drug shortage in the prior 6 months and 41% reporting shortages of more than 21 medicines during the same time frame [3].

This crisis has hit a critical juncture, with shortages leading to loss of patient lives and gross overpricing. The US Institute of Safe Medication Practices has estimated that 15 patients have died in the span of a mere 15 months because of these drug shortages [8,9]. Beyond patient mortality, hospitals have also been forced to delay treatment or leave patients untreated, in unnecessary pain, or with suboptimal clinical care, leading to complications, adverse events, and increased overall burden and costs for the health care system [3,7].

The dearth of access to these key drugs has resulted in their being sought outside the traditional drug supply chain at virtually any price and condition. Indeed, Institute of Safe Medication Practices notes that hospitals have turned to the risky secondary or "gray" market (secondary resellers outside of direct drug distribution channels) in an effort to address the drug supply concern [10]. Those entering into these markets have faced markups as high as 4500%, with *average* markups of 650% and virtually all drugs at least double the normal market price [9,10]. Despite knowledge of risks to drug safety, including counterfeiting, diversion, and substandard and falsified drugs in this market, 12% of hospitals admitted to engaging in these purchases [8]. Documented fakes entering through the gray market include counterfeit cancer, human immunodeficiency virus, and diabetes drugs [8,11,12].

However, of greatest concern is that, due to publicized profiteering, documented safety concerns, and diminishing availability of drugs in the gray market, patients affected by ongoing shortages may turn to a rapidly growing source: the Internet. However, the sale of drugs by illicit online pharmacies is of great concern, as they are a conduit for questionable therapeutic products for a wide range of treatments and diseases [13].

Indeed, purchasing drugs online is highly risky [14]. There is a significant criminal element in online drug sales, and the National Association of Boards of Pharmacy (NABP) has reported in its analysis that virtually all (96%) online drug sellers violate safe practices and/or pharmacy laws [15]. Further, there have been documented deaths from drugs obtained online, and the FDA, major national drug regulatory agencies, law enforcement, and the World Health Organization all warn against purchasing pharmaceuticals over the Internet [16-18].

Yet the challenges of drug supply shortages and exorbitant costs may drive patients in great need to seek drugs online. Given this potential crisis, we examined the availability of drugs subject to shortages on the Internet. We investigated online vendors of identified FDA current drugs in shortage to determine the characteristics of these sellers. Specifically, we examined the prevalence of online drug seller sites and intermediary sites including those sites that sell shortage drugs direct to patients, sites that act as intermediaries for the purchase of drugs and market online drug seller sites, and user-generated content via forums and social media that link to online drug sellers. We also determined whether sellers were accredited by the NABP's Verified Internet Pharmacy Practice Sites (VIPPS) accreditation program, the only accredited online pharmacies recommended by the FDA [19]. This included determination of whether vendors were on the NABP Not Recommended Sites list. Not Recommended sites are Internet drug sellers that appear out of compliance with state and federal laws or NABP patient safety and pharmacy practice standards [20]. Finally, we also assessed characteristics that could have an impact on availability or legitimacy of product, such as "international" status, which have been identified as a suspect characteristic [21]. We did not compare prices or evaluate purchasing requirements, including validating whether sellers required prescriptions, because of ethical and legal concerns arising from fictitious claims of being a patient with a particular disease, and because recent work found that online drug sellers that purportedly "require" prescriptions in fact do not require them, use "surveys" in lieu of prescriptions, and/or illicitly sell prescriptions to purchasers [14,22]. Hence, the potential quality and safety of identified sites can be better determined by examining VIPPS accreditation or Not Recommended status as outlined above. These characteristics are also important in the context of the limited ability of consumers to identify suspect drug-selling websites [23].

We also did not assess location of website registration or purported sourcing (other than as international or domestic), as these have been found to be forensically difficult to rely upon for assessment of risk characteristics [14]. For example, websites listed in Canada have been found forwarding drug orders to Israel and then financial information to Russia, where the credit card transaction was processed, and the pharmaceutical was shipped from India to a consumer in the United States. Another operated a drug-selling scheme that had its domain name hosted in Korea and registered in St Kitts, with orders dispatched from Oklahoma City [14,24,25]. Other even more dramatic examples include a seller with a Web address in Russia, server in China, payee for the credit card charge in the United Kingdom, payment processing in Australia, and product mailed from Chicago, using a return address of an unsuspecting customer of the drug website; and an Internet drug seller in Costa Rica, with computer servers located in Cyprus, credit card payments processed in Israel, and revenues placed back in bank accounts in Cyprus [26]. As well, many online drug seller websites are mirror or affiliate sites—that is, duplicate selling pages that do not sell the product itself, but instead provide the main website with greater Web presence, with the mirror or affiliate site obtaining a percentage of sales [14]. There is no requirement generally for mirror sites to be located in the same geographic area as the main website.

Methods

We first identified the FDA current shortages list of drugs with an endpoint of September 23, 2011, which included 72 drugs (*shortage drugs*) [27]. For each identified shortage drug, we searched Google using the search term *buy [drug]* where *drug* was the name of the specific shortage drug. Standard security settings were used in the browser, and no user was signed into the browser settings. We then identified the first 5 offers for

sale or, if there were fewer than 5 sales offers, the number of sales offers within the first 5 search result pages. We used 5 sales offers as our number of hits to examine based on previous research that found that consumers purchasing goods online visit 3–5 websites prior to purchase [28,29]. For this analysis, we excluded wholesalers and sponsored links. Searches were performed from October 1, 2011 to October 17, 2011.

Internet sales offers were characterized in two categories. The first identified websites as *online drug sellers* (direct vendors of the shortage drug selling direct to consumers). Online drug sellers are sometimes called Internet pharmacies or online pharmacies, but since many if not most are in fact not pharmacies [15,20], we have not adopted this term. The second category identified *intermediary sites* (websites that did not engage in the direct sale of drugs online, but acted as intermediaries or sources of information to purchase shortage drugs online). These included data aggregators (sites collecting data links for a particular shortage drug that links to online drug sellers); forum links (Internet discussion forums with links to online drug sellers); personal page data links (PPDLs, pages of personally placed links, materials, and other information that links to online drug sellers); and other social media links (eg, Facebook or Twitter).

For each identified online drug seller selling a shortage drug, we assessed whether the site was accredited by the NABP VIPPS program through review of the NABP VIPPS program website. We also assessed whether the online pharmacy was on the NABP Not Recommended Sites list, similarly reviewing the NABP Not Recommended Sites list on the NABP website [20]. We further assessed whether these online pharmacies were internationally or domestically based, regardless of their actual presence or ultimate location, as “international” status is a widely recognized characteristic of potential safety and quality concerns [21].

Finally, for intermediary sites, we assessed whether links advertised “no-prescription” drugs, an inherent indication of patient and drug safety risk as well as a violation of US law. Note that we did not include the term *no prescription* when searching for shortage drug advertisers. Non-VIPPS-accredited, international online drug sellers and intermediary sites linking to sites with no-prescription sales were deemed suspect sites.

Results

We identified 72 shortage drugs (Table 1). When searched, 68 (94%) of these drugs were being advertised online. For these 68 shortage drugs, 291 total offers for sale resulted from online searches, the vast majority (n = 207, 71.1%) by online drug sellers selling direct to consumers. Intermediary sites advertising sales of drugs in shortage included data aggregator sources (n = 22, 8%), forum links (n = 23, 8%), and PPDLs (n = 34, 12%). We also identified Flickr social media links (n = 5, 2%). All intermediary sites and Flickr social media links sales offers (n = 84, 29%) advertised drug sales without a prescription. In total, there were 91 individually identified online drug sellers, 17 forum link sources, 11 Data Aggregators, 16 PPDL sources, and 5 Flickr links. These were coded using alphanumeric designators and are listed in Table 1.

Most (n = 53, 58%) of the online drug sellers appeared to be of international origin, with Canada being the top purported source of these vendors (n = 32, 60%). We also found one online drug seller (online drug seller CB, Table 2) using the VIPPS accreditation seal, but the seller was not found on the VIPPS accreditation list. We also found one forum link also using an unauthorized VIPPS seal (see Figure 1).

Several online drug sellers dominated offers for shortage drugs (Table 2). Of the 91 vendors, 31 (34%) offered multiple (>1) drugs in shortage for sale, and were overrepresented with 148 (71%) of all online drug seller marketing events. Of these 31 sellers, 21 (68%) were on the NABP Not Recommended Sites list; 8 (26%) had no information available; only 1 (3%) was VIPPS accredited.

Finally, of the 68 drugs that were being offered for sale online, 58 (85%) were advertised for sale by at least one online drug seller (Table 1). Among these 58 drugs with online drug sellers, 53 (91%) were offered for sale by at least one Not Recommended NABP site. Indeed, 21 (36%) of these shortage drugs were advertised for sale *only* by NABP Not Recommended sites (Table 3). Only 4 drugs (7%) had any sales offers from a VIPPS-accredited pharmacy. We also found one shortage drug being advertised as over-the-counter when it is not (calcitriol; Figure 2). We also found other nonshortage drugs to be advertised as over-the-counter when they were not (eg, insulin or vaccines), as well as medical devices (eg, intrauterine devices).

Table 1. US Food and Drug Administration Current Drug Shortages list and Internet sellers (September 23, 2011)

Name of drug	Sellers
Acetylcysteine Inhalation Solution	Online drug sellers A[a], B[a], C[a], D, E
Alcohol Dehydrated (Ethanol >98%)	Online drug sellers F, G, H
Amikacin Injection	Online drug sellers I, J[a], L; forum link A
Amino Acid Injection	Data aggregator 1; forum links B, C
Aminocaproic Acid	Online drug sellers M[a], N[a], O[a], P[a]
Ammonium Chloride Injection	Online drug seller Q[a]
Ammonium molybdate injection	No sales offers
Ammonul Injection 10%/10%	Online drug seller Q[a]; forum link D; PPDLs[b] 1 (2×), 4
Amphetamine Mixed Salts, ER Capsules	Online drug sellers R, CL, S, T; PPDL 2
Anadrol-50 tablets (Oxymetholone Tablets)	Online drug sellers U, V, W[a], X, Y
Aquasol A, 50,000 units/mL, 2 mL ampule	PPDL 1; online drug seller Q[a]; forum link E
Avalide	Online drug sellers Z[a], AA, AB, A[a], AD
Bleomycin Injection	Online drug sellers AE, AI; forum links F, H, I
Buprenorphine injection	PPDLs 1, 16; online drug seller AG[a]; data aggregators 2, 3
Calcitriol 1 mcg/mL Injection	Flickr 1; PPDL 3 (2×); online drug sellers Q[a], C[a]
Calcium Chloride Injection	No sales offers
Calcium Gluconate 100 mg/mL	Online drug sellers AH[c], AI, AL, C[a]
Cerezyme (imiglucerase for injection)	Online drug seller C[a], Q[a], AJ[a]; data aggregators 4, 5
Cisplatin injection 1 mg/mL solution	Online drug sellers L[a], AK, AE; PPDL 4; data aggregator 5
Cyanocobalamin injection	Online drug sellers A[a], AM, E, AN[a], AO[a]
Cytarabine Injection	Online drug sellers AP[a], AC[a], AQ[a], AR[a], AS[a]
Daunorubicin hydrochloride solution for injection	Online drug sellers CK[a], Q[a], J[a], K; data aggregator 6
Desmopressin Injection	Online drug sellers A[a], AT[a], AU[a], AE; data aggregator 5
Dexamethasone Injection	Online drug sellers AV, J[a], K, AW; PPDL 5
Digoxin Injection	Online drug sellers AJ[a], AX; PPDLs 1, 6; data aggregator 7
Diltiazem Injection	PPDLs 1, 7; online drug sellers AY[a], AK[a], AZ[a]
Doxorubicin (adriamycin) lyophilized powder	Online drug sellers A[a], BA, BB[a], AH[c]
Doxorubicin Liposomal (Doxil) Injection	PPDL 1 (2×); online drug sellers Q[a], AG; data aggregator 8
Doxorubicin Solution for Injection	Online drug seller BC, BD, BE[a], L
Ethiodol (ETHIODIZED OIL) ampules	Online drug sellers Q[a], BF; forum link F; PPDLs 4, 8
Etoposide solution for injection	Online drug sellers C[a], A[a]; PPDL 9; forum link A
Fabrazyme (agalsidase beta)	Online drug sellers Q[a], AJ[a]; data aggregator 4; PPDL 1
Fluorouracil Injection	Online drug sellers AS, L, AY[a], AZ[a], BH
Foscarnet Sodium Injection	Online drug sellers C[a], Q[a]; data aggregators 4, 6; forum link G
Fosphenytoin Sodium Injection	Online drug sellers AS, Q[a], AY[a], AZ[a]; data aggregator 4
Furosemide Injection	Online drug sellers AP[a], AJ[a]; PPDL 9; data aggregator 9; Flickr 2

Name of drug	Sellers
Haloperidol Decanoate Injection	Online drug sellers BI[a], AH[c]; Flickr 3; PPDL 1
Intravenous Fat Emulsion	Online drug sellers Q[a]; data aggregators 6, 5; PPDL 1; forum link H
Isoniazid Tablets	Online drug sellers AO[a], BJ[a], D; forum links I, A
Leucovorin Calcium Lyophilized Powder for Injection	Online drug sellers C[a], AG[a], AJ[a], K; forum link J
Leuprolide Injection	Online drug sellers AT[a], A[a], BJ[a], P[a], AU[a]
Levoleucovorin (Fusilev) 50 mg single use vials	Online drug sellers Q[a], BG; PPDL 1
Lorazepam Injection	Online drug seller BK
Magnesium Sulfate Injection	Online drug sellers BL[a], J[a], K, L, AY[a]
Methylphenidate HCl	Online drug sellers BM[a], AG[a], BN; forum links K, L
Metoclopramide injection	Online drug seller AJ[a]; forum links M, N, O; Flickr 4
Mitomycin Powder for Injection	Online drug sellers BO[a], BP[a], BQ, BE[a], BR
Multi-Vitamin Infusion (Adult and pediatric)	PPDL 11; data aggregator 5
Nalbuphine Injection	Online drug sellers BS, BM[a], BT; forum link P; PPDL 12
NeoProfen (ibuprofen lysine) Injection	Online drug sellers AJ[a], AX[a], BU; PPDL 13; forum link P
Neostigmine methylsulfate injection	Online drug sellers J[a], AG[a], K, AT; data aggregator 10
Neupro (rotigotine transdermal system)	Online drug sellers A[a], Q[a], AT[a], AJ[a], AO[a]
Norepinephrine Bitartrate Injection	Online drug sellers C[a], Q[a]; data aggregator 6; forum links G, H
Ontak injection	Online drug seller Q[a]; data aggregators 4, 5; forum link G; PPDL 1
Oxsoralen (methoxsalen) 1% topical lotion	Online drug sellers AG[a], AO[a]; data aggregator 10; forum link Q
Oxsoralen-Ultra (methoxsalen) 10 mg capsules	Online drug sellers Q[a], AX[a], BV, BW, AO[a]
Paclitaxel Injection	Online drug sellers BI[a], BX, L, BY, D
Phenylephrine HCl Injection	PPDLs 14, 15
Potassium Phosphate Injection	No sales offers
Procainamide HCL Injection	Online drug sellers BZ, CA[a], A[a]
Propofol Injection	Online drug sellers A[a], C[a], AT[a], Q[a], BB[a]
Sodium Chloride 23.4%	Online drug sellers Q[a], AI, CC, CD, CB[d]
Sodium Chloride 14.6% Injection	Online drug seller AH[c]
Sodium Phosphate Injection	No sales offers
Streptomycin for Injection, USP	Online drug sellers CF[a], BE[a], Q[a], CE
Sulfamethoxazole 80mg/trimethoprim 16mg/ml injection (SMX/TMP)	Online drug sellers Q[a], AJ[a], AX[a], CG; data aggregator 11
Thiotepa for Injection	Data aggregators 4, 5
Thyrogen (thyrotropin alfa) injection 1.1mg/vial	Online drug sellers AJ[a], CH[a], Q[a]; data aggregator 4; PPDL 4
Thyrolar Tablets	Data aggregator 4
Vasopressin Injection	Online drug sellers J[a], AS[a], K, CI, CJ[a]
Vecuronium Injection	Online drug sellers Q[a], AS, CM, AY; forum link G

Name of drug	Sellers
Vincristine Sulfate Injection	Online drug sellers A[a], AS, CM, L; PPDL 1

[a] On National Association of Boards of Pharmacy (NAPB) Not Recommended Sites list for online drug purchasing.

[b] Personal page data link.

[c] NAPB Verified Internet Pharmacy Practice Sites (VIPPS)-accredited site.

[d] Claims NABP VIPPS accreditation but not on NABP VIPPS accreditation list.

Figure 1. Unauthorized uses of Verified Internet Pharmacy Practice Sites (VIPPS) seal.

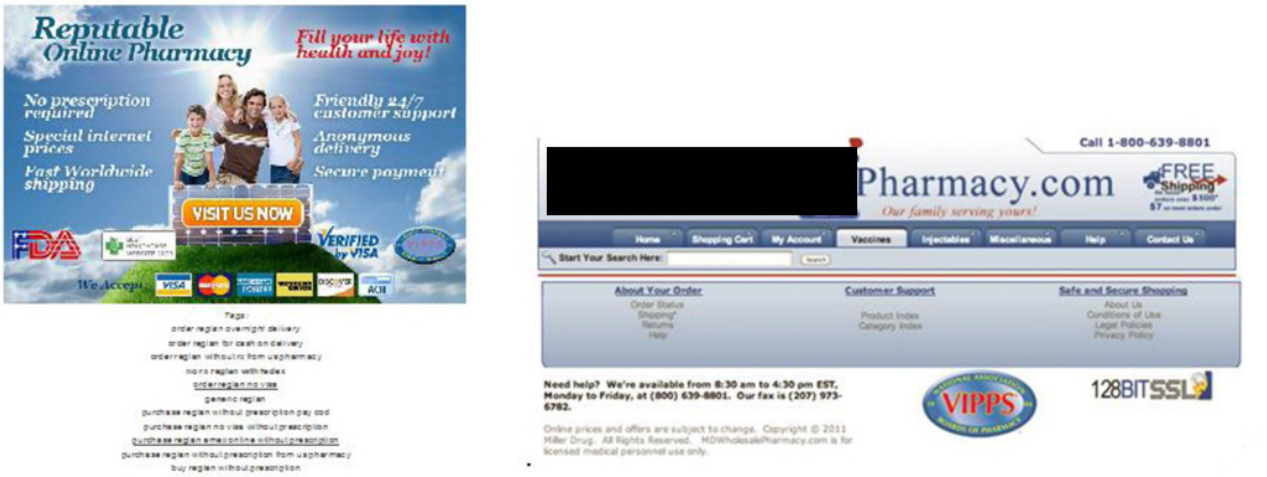


Figure 2. Shortage drug calcitriol injection advertised as an over-the-counter drug.

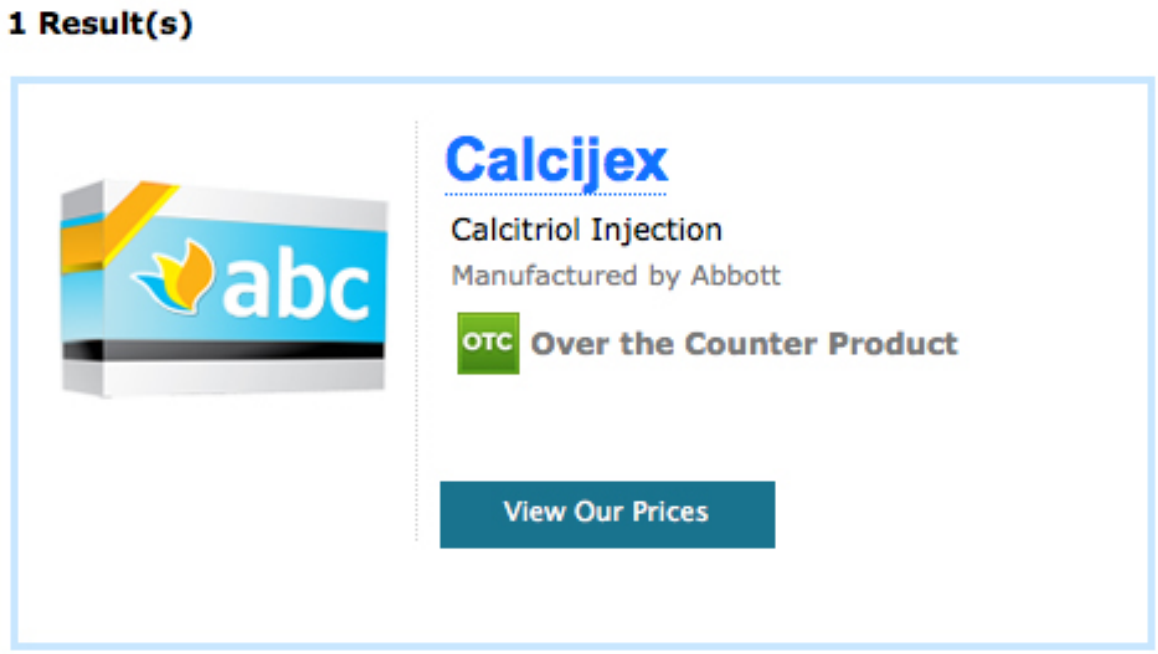


Table 2. Characteristics of Internet shortage drug sellers

Online drug seller	Marketing event frequency	NABP[a] status	International/domestic (US) status
Q	23	Not recommended	International
A	12	Not recommended	International[b]
AJ	10	Not recommended	International[b]
C	9	Not recommended	International
K	7	No information	
L	7	No information	International
J	6	Not recommended	International[b]
AG	6	Not recommended	
AS	6	Not recommended	
AO	5	Not recommended	International[b]
AT	5	Not recommended	International[b]
AY	5	Not recommended	International[b]
AH	4	VIPPS[c] accredited	US
AX	4	Not recommended	
AE	3	No information	International[b]
BJ	3	Not recommended	International[b]
D	3	No information	International[b]
AZ	3	Not recommended	International[b]
BE	3	Not recommended	International
P	2	Not recommended	International[b]
AI	2	No information	
AK	2	Not recommended	International[b]
AP	2	Not recommended	International
AU	2	Not recommended	International
BB	2	Not recommended	International[b]
BI	2	Not recommended	International[b]
BM	2	Not recommended	International[b]
CM	2	No information	International[b]
E	2	No information	International[b]
AP	2	Not recommended	
BG	2	No information	
N	1	Not recommended	International[b]
M	1	Not recommended	International[b]
B	1	Not recommended	International[b]
F	1	No information	
G	1	No information	
H	1	No information	
I	1	No information	International

Online drug seller	Marketing event frequency	NABP[a] status	International/domestic (US) status
O	1	Not recommended	International
R	1	No information	
S	1	No information	International
T	1	No information	
U	1	No information	
V	1	No information	
W	1	Not recommended	
X	1	No information	
Y	1	No information	International
AA	1	No information	
AB	1	No information	International
AD	1	No information	
AF	1	No information	
AL	1	No information	
AM	1	Not recommended	
AN	1	Not recommended	International[b]
AQ	1	Not recommended	International
AR	1	Not recommended	
AV	1	No information	
AW	1	No information	International[b]
BA	1	No information	International
BC	1	No information	International
BD	1	No information	International
BF	1	No information	
BH	1	VIPPS accredited	
BK	1	No information	
BL	1	Not recommended	International
BN	1	No information	International
BO	1	Not recommended	
BP	1	Not recommended	International[b]
BQ	1	No information	International[b]
BR	1	No information	International
BS	1	No information	
BT	1	No information	
BU	1	No information	International
BV	1	No information	International
BW	1	No information	International[b]
BX	1	No information	
BY	1	No information	International
BZ	1	No information	
CA	1	No information	International[b]

Online drug seller	Marketing event frequency	NABP[a] status	International/domestic (US) status
CB[d]	1	Not recommended	
CC	1	No information	
CD	1	No information	
CE	1	No information	
CF	1	Not recommended	International[b]
CG	1	No information	International[b]
CH	1	Not recommended	International[b]
CI	1	No information	
CJ	1	Not recommended	International[b]
CK	1	Not recommended	International
CL	1	No information	
Z	1	Not recommended	International[b]

[a] National Association of Boards of Pharmacy.

[b] Purported Canadian websites.

[c] Verified Internet Pharmacy Practice Sites.

[d] Use of VIPPS seal by nonaccredited vendor.

Table 3. US Food and Drug Administration Current Drug Shortage list drugs found only on National Association of Boards of Pharmacy's Not Recommended online drug seller sites

Drug
Aminocaproic acid
Ammonul
Calcitriol
Cerezyme
Cytarabine
Diltiazem
Doxorubicin
Etoposide solution
Fabrazyme
Foscarnet
Furosemide
Intravenous fat emulsion
Leuprolide
Magnesium sulfate injection
Metoclopramide
Neupro
Norepinephrine injection
Ontak injection
Oxsoralen 1% topical
Propofol
Thyrogen injection

Discussion

Patients face considerable risks when attempting to procure FDA current shortage drugs online. We find that suspect vendors populate much of the online market for these drugs, and international online drug sellers identified by NABP as Not Recommended dominate this eHealth landscape. With more than 90% of these drugs being offered for sale by at least one NABP Not Recommended site, in effect if an online drug seller is selling an FDA shortage drug, there is virtually always an NABP Not Recommended site selling in that market. Indeed, with more than a third of all of these drugs, including cancer, emergency, orphan drug, and other lifesaving therapeutic products, being sold by *only* NABP Not Recommended site vendors, it is highly likely that anyone searching online for drugs in shortage will encounter a suspect seller. Even those patients who may seek to legally procure a shortage drug online with a valid prescription may have a difficult time in identifying a legitimate online drug seller source, given oversaturation of suspect providers and marketing sources. This is of great patient safety concern, particularly in the context that most online drug sellers are suspect and research indicating that online drug sellers magnify positive aspects of online purchases but minimize discussion of associated risks [30].

In addition, with 100% of the intermediary sites marketing these scarce medications as no-prescription drugs (see Figure 3), including social media sites such as Flickr that had been previously unobserved (see Figure 4) [29], this further indicates that the online marketplace for these treatments is extremely risky and diversifying.

These findings, in combination with previous work showing the extensive presence of illicit online drug sellers in other social media such as Facebook and Twitter [31,32], should place all stakeholders on alert about the legitimacy of online drug sellers, other vendors, and the quality of their products.

It is particularly troubling that so few NABP VIPPS-accredited pharmacies appeared in the returns from a common Google search. Even if legitimate pharmacies are marketing shortage drugs, the overwhelming presence of NABP Not Recommended online drug sellers may crowd out their arising early or at all in search results. The tremendous dominance of suspect online drug sellers and the virtual absence of authorized VIPPS-accredited online pharmacies for drugs in short supply may have an immediate impact on patient safety. This is particularly true in the context of other work reporting similar findings for online availability of biologics in short supply, which also include absence of VIPPS-accredited vendors, as well as unauthorized use of the VIPPS seal and online drug seller marketing of shortage vaccines as over-the-counter drugs [31].

In addition, drug shortages, documented and publicized safety risks, and sales profiteering by gray market wholesalers not only may lead patients to attempt to procure shortage drugs through the Internet, but also may lead some providers to seek out shortage drug treatments from these online vendors. Not only is online purchasing from these vendors inherently dangerous, with compromised and counterfeit medications injuring and killing unsuspecting patients globally [13], but also scarcity in the legitimate market questions the very premise that these online drug sellers are selling authentic medication. Given current market-based demands and shortages, it does not appear that online drug sellers would have such an abundant supply of shortage drugs as advertised, and health care providers should regard these offers as highly suspect.

Indeed, even if these shortage drugs were authentic, Not Recommended online drug sellers, particularly international-based sites highly prevalent in this study, would be unlikely to have adequate knowledge and impetus to ensure proper transport [8]. Many of these drugs, including biologics, are sensitive and require special temperature controls and handling to retain therapeutic effectiveness—quality concerns that may not be known or of concern to suspect vendors [10,31]. Previous reports have also noted issues with online drug sellers' storage and transport of sensitive drugs [33,34].

The dominance of international online sellers should be of particular concern, as the FDA does not permit personal drug importation from international sites due to its lack of ability to control quality and oversight of foreign materials [35]. Importing these drugs through purchase from these online drug sellers consequently violates the US Food, Drug, and Cosmetic Act [35]. Further, an argument for importation under an FDA personal importation exemption is inapplicable, even for the vast majority of these drugs, since it generally requires both that the drug be commercially unavailable in the United States and that use of any approved importation be medically supervised [36].

Moreover, it should also be noted that, from the patient's perspective, drugs purchased from nondomestic sources are not eligible for public program reimbursement. Such purchases may lead to unnecessary patient expenditures, including remediation expenditures for additional care [36].

This online environment with its limited regulation and cloaked nature as reviewed by Orizio et al [37], combined with the current drug shortage, is placing tremendous strain on the safety of the drug supply and fueling illicit activity. This situation exploits the desperate nature of vulnerable patients seeking any treatments that could potentially influence the course of their disease, similar to the Laetrile frauds for cancer treatments from the 1970s [26,38]. With the addition of reports that hospitals are increasingly sourcing drugs in shortage from the gray market [8,9], providers and patients seeking these therapies are surrounded by safety threats.

In response, at a minimum, risk management approaches should be considered. The US Customs and Border Protection agency is overwhelmed in attempting to monitor incoming drugs. It cannot destroy contraband suspect drugs coming through the mails due to international postal conventions. In general, it is limited to returning the package and contents to the original sender (for potential illicit resale) if not assessed within 24 hours, or simply allowing it to be sent on for delivery[14,39-41]. Hence, demand-side actions should be a focus.

Figure 3. Intermediary site marketing products as no-prescription drug.

Figure 4. Flickr social media link for no-prescription shortage drug.

Providers seeking drugs in short supply should verify the authenticity of the drug pedigree (ie, its documentation of transport) and should take advantage of professional resources for safe sourcing of drugs [10,42]. This includes implementing facility safety and regulatory protocols ensuring drug sourcing through accredited wholesalers and vendors that can and do certify that they procure drugs from legitimate, accredited distributors in compliance with pharmaceutical distributor laws, licensure, and regulations. NABP also provides accreditation here, and providers should purchase only from these NABP-accredited vendors [43].

This approach of listing and updating legitimate wholesale sources of drugs by a neutral, trusted entity should be considered as a strategy more broadly to allow rapid determination of vendor legitimacy on regional levels globally. These regional entities can then share information as to identified problematic vendors for information coordination, as well as determining and communicating best practices in promoting safety measures in the drug supply chain.

Patients should also be counseled about the dangers of counterfeit and substandard drugs when purchasing drugs online, and be advised to purchase only from verified NABP VIPPS-accredited online drug sellers as recommended by the FDA [16]. They should also be encouraged to inquire about provider drug-procurement policies, and require that drugs administered to them have appropriate pedigree evidencing authenticity. The patient is the last barrier to harm and should be partnered with and actively engaged in patient safety efforts. Such an approach is amenable to application in other health-delivery environments as well, with consistent public health messages and patient education such as that promulgated by important global health groups such as the International Council of Nurses [44]. These public health messages should expressly note global drug regulatory authority warnings and updated information on the dangers of purchasing drugs—including shortage drugs—online [18,19].

These drug shortages have historically been a problem in clinical care, and have been the subject of past calls for reform and attention due to their adverse impact on public health [10,45-48]. Indeed, drug shortages have tripled in the last 6 years, while few have left the list, portraying a worrying trend [3], while the global nature of the problem has become increasingly dire [49,50].

Beyond reactive approaches, organized, proactive planning by stakeholders to address this issue should be started as soon as possible. Existing calls for enhanced drug regulatory authority powers to address drug shortages, stronger regulation of pharmaceutical distribution and pricing incentives, and more stringent penalties against illegal profiteering activities should be supported by stakeholders' joining together in public–private partnership models to plan for new challenges in pharmaceutical supplies and safety.

These public–private partnerships should include multisector engagement with pharmaceutical manufacturers (including both brand and generic industries), drug regulators, patient safety and advocacy groups, hospitals, group purchasing organizations, and professional societies to prioritize and develop organized responses to drug shortages. Domestic and regional best practices can be applied in proactive efforts as well, with shared systems planning and strategies among stakeholders integrated across geopolitical and health-delivery environments. At the outset, a risk management approach may be best for priority planning—for example, public–private partnerships organized around supplying chemotherapy drugs and emergency department treatments, whose absence may have the greatest potential adverse clinical impact. In addition, due to the security threat illegal pharmaceutical sales pose and the need to address drug shortages as a national emergency, national security agencies and legal authorities should also be included in these public–private partnerships.

Integrated industry, regulator, patient, and provider partnerships would be able to better anticipate potential shortages and their impact on patient safety, and to help develop information resources and responses for stakeholders that are currently lacking for these situations [51]. Further, these partnerships would enhance existing calls for more advanced notification of shortages by manufacturers through establishing lines of communication [3,47,52,53] and would help to develop market-based incentives to encourage investment in manufacturing of shortage drugs, such as reimbursement-rate adjustment and creative financing for increased production [6,54].

Finally, strategic plans for safe substitutions of shortage drugs that could be integrated into emergency planning for regional delivery systems, along with voluntary international coordination of emergency supplies, should be considered in particularly needed areas and populations (such as public health treatments for pandemics or other communicable disease events). These partnerships would be similar to local coalition efforts to reallocate drugs subject to shortages by borrowing, back-ordering, increasing surveillance and legal enforcement against potential counterfeits, and seeking safe alternative supply sources through greater regional coordination [8,33].

We note that this study has several limitations. It evaluates the drugs on the FDA Current Drug Shortages list as of one point in time, and its assessment of online availability similarly is limited to that period. Further, websites arise, change, and are taken down dynamically on the Internet, so the information and websites found in this study are necessarily limited as well. NABP VIPPS-accredited online pharmacies and NABP Not Recommended site designations also are limited in validity over time due to changes in the Internet and vendors. Finally, we did not determine the actual quality of products through test purchasing, as buying drugs for a fictional patient raises ethical and legal concerns.

Overall, drugs on the FDA Current Drug Shortages list are critical therapeutic tools in the medical arsenal. Because of the crisis of access to these drugs, sources such as the Internet may be considered convenient for procurement. However, suspect vendors appear to dominate online marketing of shortage drugs. This high-risk digital conduit of medicines should be addressed. Reactive and proactive risk management strategies should be engaged to ensure that providers render optimal, safe care to patients, and that patients who are considering purchasing online

due to desperate circumstances do not become the next victims of the drug shortage. Extreme caution and caveat emptor may best describe the themes when approaching this eHealth market.

Acknowledgments

BAL presented this work in part at the Global Drug Safety Roundtable, Council on Foreign Relations, December 5, 2011, and thanks that audience for useful comments. TKM is the 2011–2012 Carl L. Alsberg, MD, Fellow of the Partnership for Safe Medicines, the latter of which is not connected with the submitted work, and gratefully acknowledges that support.

Conflicts of Interest

None declared.

References

1. Kaiser J. Medicine. Shortages of cancer drugs put patients, trials at risk. Science 2011 Apr 29;332(6029):523. [doi: 10.1126/science.332.6029.523] [Medline: 21527686]
2. Emanuel EJ. The New York Times. 2011 Aug 6. Shortchanging Cancer Patients URL: http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en/us/doubleclick/pdfs/DoubleClick-02-2005-Search-Before-the-Purchase.pdf [accessed 2011-12-16] [WebCite Cache ID 63zfKUKvM]
3. Mitka M. FDA, US hospital and pharmacy groups report drug shortages a growing problem. JAMA 2011 Sep 14;306(10):1069-1070. [doi: 10.1001/jama.2011.1238] [Medline: 21917568]
4. Anonymous. Shortages of cancer drugs in the USA. Lancet Oncol 2011 Apr;12(4):313. [doi: 10.1016/S1470-2045(11)70087-0] [Medline: 21463830]
5. Dolgin E. NIH faces marching orders on orphan drug shortage. Nat Med 2011 May;17(5):522. [doi: 10.1038/nm0511-522b] [Medline: 21546953]
6. Roehr B. FDA is to hold public meeting to discuss reasons for drug shortages. BMJ 2011;343:d5650. [Medline: 21893564]
7. Cherici C, Frazier J, Feldman M, Gordon B, Petrykiw C, Russel W, et al. Premiere Inc. 2011 Mar. Navigating Drug Shortages in American Healthcare: A Premier Healthcare Alliance Analysis URL: http://www.premierinc.com/about/news/11-mar/drug-shortage-white-paper-3-28-11.pdf [accessed 2011-11-16] [WebCite Cache ID 63FEdqqFL]
8. Institute for Safe Medication Practices. 2011 Aug 25. Gray Market, Black Heart: Pharmaceutical Gray Market Finds a Disturbing Niche During the Drug Shortage Crisis URL: http://www.ismp.org/Newsletters/acutecare/showarticle.asp?ID=3 [accessed 2011-11-16] [WebCite Cache ID 63FEeP0jZ]
9. Johnson LA. Yahoo! News. 2011 Sep 24. AP IMPACT: Hospital Drug Shortages Deadly, Costly. URL: http://news.yahoo.com/ap-impact-hospital-drug-shortages-deadly-costly-113347849.html [accessed 2011-11-16] [WebCite Cache ID 63FEm40ak]
10. Cherici C, McGinnis P, Russell W. Premier Inc. 2011 Aug. Buyer Beware: Drug Shortages and the Gray Market URL: http://www.premierinc.com/about/news/11-aug/Gray-Market/Gray-Market-Analysis-08152011.pdf [accessed 2011-11-16] [WebCite Cache ID 63FFqqO0j]
11. Kane J. PBS Newshour. 2011 Aug 29. Drug Prices Soar as Hospital Suppliers are Forced into 'Gray Market' URL: http://www.pbs.org/newshour/rundown/2011/08/drug-prices-soar-as-pharmacists-are-forced-into-gray-market.html [accessed 2011-11-16] [WebCite Cache ID 63FFWTFBG]
12. Clarke PE. US Food and Drug Administration. 2011 Jun 9. FDA Working to Lessen Patient Impact from Drug Shortages URL: http://www.fda.gov/Drugs/DrugSafety/DrugShortages/ucm257745.htm [accessed 2011-11-16] [WebCite Cache ID 63FEvm1kL]
13. Mackey TK, Liang BA. The global counterfeit drug trade: patient safety and public health risks. J Pharm Sci 2011 Nov;100(11):4571-4579. [doi: 10.1002/jps.22679] [Medline: 21698604]
14. Liang BA, Mackey T. Searching for safety: addressing search engine, website, and provider accountability for illicit online drug sales. Am J Law Med 2009;35(1):125-184. [Medline: 19534258]
15. National Association of Boards of Pharmacy. 2011 Jul 28. NABP Issues Rogue Online Pharmacy Public Health Alert URL: http://www.nabp.net/news/nabp-issues-rogue-online-pharmacy-public-health-alert/ [accessed 2011-11-16] [WebCite Cache ID 63FEwAfAz]
16. US Food and Drug Administration. 2011 Oct 19. The Possible Dangers of Buying Medicines over the Internet URL: http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm048396.htm [accessed 2011-11-16] [WebCite Cache ID 63FF1isHJ]
17. US Department of Justice, Drug Enforcement Administration, Office of Diversion Control. Consumer Alert: Prescription Drugs: Buying On-line Could Mean Doing Time URL: http://www.deadiversion.usdoj.gov/consumer_alert.htm [accessed 2011-11-16] [WebCite Cache ID 63FF1YGlj]
18. World Health Organization. 2010 Jan. Medicines: Spurious/Falsely-Labelled/Falsified/Counterfeit (SFFC) Medicines URL: http://www.who.int/mediacentre/factsheets/fs275/en/ [accessed 2011-11-16] [WebCite Cache ID 63FFAWYkc]

19. US Food and Drug Administration. What Can I Do to Help Ensure I Do Not Receive or Take Counterfeit Medicine? 2010 Dec 14 URL: http://www.fda.gov/AboutFDA/Transparency/Basics/ucm235684.htm [accessed 2011-11-16] [WebCite Cache ID 63FFKfMSC]
20. National Association of Boards of Pharmacy. 2012. Not Recommended Sites URL: http://www.nabp.net/programs/consumer-protection/buying-medicine-online/not-recommended-sites/ [accessed 2011-11-16] [WebCite Cache ID 63FFbPyEt]
21. National Association of Boards of Pharmacy. 2012. VIPPS: Set Yourself Apart From Rogue Sites URL: http://www.nabp.net/programs/accreditation/vipps/ [accessed 2011-11-16] [WebCite Cache ID 63FFaxGt6]
22. Griffin G. CNN. 2009 Feb 2. CNN Special Investigation Report: Illegal Online Pharmacies URL: http://www.cnn.com/video/ [accessed 2011-12-16] [WebCite Cache ID 63zfbv5JX]
23. Orizio G, Schulz P, Domenighini S, Bressanelli M, Rubinelli S, Caimi L, et al. Online consultations in cyberpharmacies: completeness and patient safety. Telemed J E Health 2009 Dec;15(10):1022-1025. [doi: 10.1089/tmj.2009.0069] [Medline: 19929235]
24. Work D. Winston-Salem Journal. 2009 Feb 7. Phony Medicines Available Online: Bogus Pharmacies on Internet Spread Drugs That Can Be Dangerous, Deadly URL: http://www2.journalnow.com/news/2009/feb/07/phony-medicines-available-online-ar-138626/ [accessed 2011-12-17] [WebCite Cache ID 63zfiVJNZ]
25. Jackson G. Faking it: the dangers of counterfeit medicine on the internet. Int J Clin Pract 2009 Feb;63(2):181. [doi: 10.1111/j.1742-1241.2008.01989.x] [Medline: 19196352]
26. Liang BA. A dose of reality: promoting access to pharmaceuticals. Wake Forest Intellect Prop Law J 2008;8(3):301-386.
27. US Food and Drug Administration. 2011 Nov 16. Current Drug Shortages URL: http://www.fda.gov/drugs/drugsafety/drugshortages/ucm050792.htm [accessed 2011-11-16] [WebCite Cache ID 63FFB6W9D]
28. Ahuja M, Gupta B, Raman P. An empirical investigation of online consumer purchasing behavior. Commun ACM 2003 Dec;46(12):145. [doi: 10.1145/953460.953494]
29. Doubleclick.com. 2005 Feb. Search Before the Purchase URL: http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en/us/doubleclick/pdfs/DoubleClick-02-2005-Search-Before-the-Purchase.pdf [accessed 2011-11-16] [WebCite Cache ID 63FFLRv70]
30. Orizio G, Rubinelli S, Schulz PJ, Domenighini S, Bressanelli M, Caimi L, et al. "Save 30% if you buy today". Online pharmacies and the enhancement of peripheral thinking in consumers. Pharmacoepidemiol Drug Saf 2010 Sep;19(9):970-976. [doi: 10.1002/pds.2007] [Medline: 20652863]
31. Liang BA, Mackey TK. Vaccine shortages and suspect online pharmacy sellers. Vaccine 2012 Jan 5;30(2):105-108. [doi: 10.1016/j.vaccine.2011.11.016] [Medline: 22094281]
32. Liang BA, Mackey TK. Prevalence and Global Health implications of social media in direct-to-consumer drug advertising. J Med Internet Res 2011;13(3):e64 [FREE Full text] [doi: 10.2196/jmir.1775] [Medline: 21880574]
33. Weiss AM. Buying prescription drugs on the internet: promises and pitfalls. Cleve Clin J Med 2006 Mar;73(3):282-288 [FREE Full text] [Medline: 16548451]
34. Orizio G, Schulz P, Domenighini S, Caimi L, Rosati C, Rubinelli S, et al. Cyberdrugs: a cross-sectional study of online pharmacies characteristics. Eur J Public Health 2009 Aug;19(4):375-377 [FREE Full text] [doi: 10.1093/eurpub/ckn146] [Medline: 19151103]
35. Liang BA. Fade to black: importation and counterfeit drugs. Am J Law Med 2006;32(2-3):279-323. [Medline: 16927613]
36. Cooperman T, Lewitt G. Searching for safety: addressing search engine, website, and provider accountability for illicit online drug sales. Am J Law Med 2009;35(4):682-3; author reply 683-90. [Medline: 20196287]
37. Orizio G, Merla A, Schulz PJ, Gelatti U. Quality of online pharmacies and websites selling prescription drugs: a systematic review. J Med Internet Res 2011;13(3):e74 [FREE Full text] [doi: 10.2196/jmir.1795] [Medline: 21965220]
38. Schulte G. An overview of pharmaceutical smuggling cases in San Diego: it goes better when agencies work together. J Biolaw Bus 2006;9(4):41-45.
39. Federal Food, Drug, and Cosmetic Act, 21 USC §381(a) (2000). URL: http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/FDCActChapterVIIIImportsandExports/ucm107032.htm [accessed 2012-01-31] [WebCite Cache ID 656h9Hr6W]
40. Anonymous. Hearing on Refusal of Admission, 21 CFR §1.94 (2008). URL: http://law.justia.com/cfr/title21/21-1.0.1.1.1.4.16.4.html [accessed 2012-01-31] [WebCite Cache ID 656hIYoXq]
41. Stana RM. US Government Accountability Office. 2005. Prescription Drugs: Enhanced Efforts and Better Agency Coordination Needed to Address Illegal Importation: Testimony Before the Subcommittee on Oversight and Investigations, 109th Cong. 33 () URL: http://www.gao.gov/new.items/d06175t.pdf [accessed 2011-12-17] [WebCite Cache ID 63zfwhFTz]
42. Anonymous. Partnership for Safe Medicines. 2011. L.E.A.D.E.R.'s Guide For Pharmacists URL: http://www.safemedicines.org/leaders-guide-for-pharmacists.html [accessed 2011-11-16] [WebCite Cache ID 63FFKvM2g]
43. National Association of Boards of Pharmacy. 2011 Nov 16. VAWD-Accredited Facilities URL: http://www.nabp.net/programs/accreditation/vawd/vawd-accredited-facilities/ [accessed 2011-11-16] [WebCite Cache ID 63FFbsYZz]
44. International Council of Nurses. 2010 May 12. 2005 Counterfeits Kill: Nurses Target Counterfeit Medicines URL: http://www.icn.ch/publications/2005-counterfeits-kills/ [accessed 2011-12-17] [WebCite Cache ID 63zg18psH]

45. Steinbrook R. Drug shortages and public health. N Engl J Med 2009 Oct 15;361(16):1525-1527. [doi: 10.1056/NEJMp0906922] [Medline: 19828529]
46. Charatan F. Prescription drug shortages plague US. BMJ 2001 Jan 20;322(7279):130. [Medline: 11159566]
47. Hampton T. Experts look for ways to lessen impact of drug shortages and discontinuations. JAMA 2007 Aug 15;298(7):727-728. [doi: 10.1001/jama.298.7.727] [Medline: 17699002]
48. Jensen V, Rappaport BA. The reality of drug shortages--the case of the injectable agent propofol. N Engl J Med 2010 Aug 26;363(9):806-807 [FREE Full text] [doi: 10.1056/NEJMp1005849] [Medline: 20554977]
49. Health Canada. 2011 Nov 21. Web info sheet on Drug Shortages URL: http://www.hc-sc.gc.ca/dhp-mps/prodpharma/activit/announce-annonce/shortage-rupture-eng.php [accessed 2011-12-17] [WebCite Cache ID 63zg4pf9K]
50. Generics and Biosimilars Initiative. 2011 Nov 18. Biosimilars and Cancer Drug Shortages in Europe URL: http://www.gabionline.net/layout/set/print/content/view/full/1555.%20Accessed%202011-11-29 [accessed 2011-12-17] [WebCite Cache ID 63zg8dUba]
51. Kaakeh R, Sweet BV, Reilly C, Bush C, DeLoach S, Higgins B, et al. Impact of drug shortages on U.S. health systems. Am J Health Syst Pharm 2011 Oct 1;68(19):1811-1819. [doi: 10.2146/ajhp110210] [Medline: 21930639]
52. Thompson CA. House legislation proposes early warning system for drug shortages. Am J Health Syst Pharm 2011 Aug 1;68(15):1379-1381. [doi: 10.2146/news110052] [Medline: 21785017]
53. Barlas S. Severe drug shortages impose heavy costs on hospital pharmacies: senate bill might help … or not. P T 2011 May;36(5):242-302. [Medline: 21785536]
54. Connelly P, Quinn BP. Manufacturing decline yields drug shortages. Science 2011 Jul 8;333(6039):156-157. [doi: 10.1126/science.333.6039.156-b] [Medline: 21737721]

Abbreviations

FDA: Food and Drug Administration
NABP: National Association of Boards of Pharmacy
PPDL: personal page data link

Edited by G Eysenbach; submitted 16.11.11; peer-reviewed by R Levaggi, I Montoya, L Ivanitskaya; comments to author 07.12.11; revised version received 19.12.11; accepted 16.01.12; published 09.02.12.

Please cite as:
Liang BA, Mackey TK
Online Availability and Safety of Drugs in Shortage: A Descriptive Study of Internet Vendor Characteristics
J Med Internet Res 2012;14(1):e27
URL: http://www.jmir.org/2012/1/e27/
doi:10.2196/jmir.1999
PMID:22321731

Letter

How to Create Memorizable and Strong Passwords

Pietro Cipresso[1,2], MSc, PhD; Andrea Gaggioli[1,2], MPsych, PhD; Silvia Serino[1,2], MPsych; Sergio Cipresso[3]; Giuseppe Riva[1,2], MA, MPsych, PhD

[1]Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Milano, Italy

[2]Psychology Department, Catholic University of Milan, Milano, Italy

[3]Freelancer, Milan, Italy

Corresponding Author:
Pietro Cipresso, MSc, PhD
Applied Technology for Neuro-Psychology Lab
IRCCS Istituto Auxologico Italiano
Via Pellizza da Volpedo 41
Milano, 20149
Italy
Phone: 39 61911 ext 2892
Fax: 39 619112892
Email: p.cipresso@auxologico.it

(J Med Internet Res 2012;14(1):e10) doi:10.2196/jmir.1906

KEYWORDS

Privacy; security; passwords; psychology

How to Create Memorizable and Strong Passwords

In a recent JMIR article, El Emam, Moreau and Jonker highlight the importance of using strong passwords to protect personal health information in clinical trials [1]. An important implication that was not fully discussed is the potential problem people may have to create passwords that are complex but at the same time easy to remember.

To address this problem we propose the PsychoPass methord, a simple way to create strong passwords which are easy to remember. This method relies on mental practice and is not an hardware or a software to download. The idea is that a password can be created, memorized and recalled by just thinking of an *action sequence* instead of a word or string of characters. To be more specific, the method consists of the following steps (see Figure 1 and 2): (1) begin with a letter on the keyboard; (2) memorize a sequence of actions (something like “the key on the left, then the upper one, then the one on the right”, and so on); (3) memorize the sequence (not the letters used); (4) create as many passwords as you want by remembering only the first letter and the sequence. Using different types of sequences it is possible generate thousands of different passwords. Using sequences' combination is possible to create an infinite number of passwords. Moreover the created passwords will be a nonsense sequence of letters, numbers and symbols, resilient to any attack.

Furthermore the password communication among colleagues maybe done just by using the first letter and on the base of a common knowledge of the sequence (e.g., sequence 3, letter j).

El Emam and Colleagues state that more sophisticated collaboration tools are required to allow file sharing without password sharing, and provide several recommendations to implement these practices. We think that more awareness and new practices among users may represent the correct way to implement security beyond the technological issues. In particular, future research needs to focus on the processes that make technology a powerful tool for security.

Figure 1. The same sequence to generate different passwords (about 15 minutes to memorize the sequence).

Figure 2. Another sequence to generate other passwords (about 15 minutes to memorize the sequence).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Movie demonstrating the method.

[MOV File, 7MB - jmir_v14i1e10_app1.MOV]

Reference

1. El Emam K, Moreau K, Jonker E. How strong are passwords used to protect personal health information in clinical trials? J Med Internet Res 2011;13(1):e18 [FREE Full text] [doi: 10.2196/jmir.1335] [Medline: 21317106]

Edited by G Eysenbach; submitted 06.08.11; peer-reviewed by K El Emam; accepted 07.01.12; published 10.01.12.

Please cite as:
Cipresso P, Gaggioli A, Serino S, Cipresso S, Riva G
How to Create Memorizable and Strong Passwords
J Med Internet Res 2012;14(1):e10
URL: http://www.jmir.org/2012/1/e10/
doi:10.2196/jmir.1906
PMID:22233980

Corrigenda and Addenda

Correction: Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact

Gunther Eysenbach[1,2,3], MD, MPH, FACMI

[1]University Health Network, Centre for Global eHealth Innovation & Techna Institute, Toronto, ON, Canada
[2]Institute for Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada
[3]JMIR Publications Inc., Toronto, ON, Canada

Corresponding Author:
Gunther Eysenbach, MD, MPH, FACMI
University Health Network
Centre for Global eHealth Innovation & Techna Institute
190 Elizabeth St
Toronto, ON, M4L3Y7
Canada
Phone: 1 416 7866970
Fax: 1 416 340 3595
Email: geysenba@uhnres.utoronto.ca

Related Article:

Correction of: http://www.jmir.org/2011/4/e123

(J Med Internet Res 2012;14(1):e7) doi:10.2196/jmir.2041

A minor error in the references section in the originally published version of the editorial by Eysenbach (J Med Internet Res 2011;13[4]:e123) on the relationship between citations and tweetations has been corrected; in addition, references being part of the dataset are no longer cited as “references”. The now corrected problem with the references was a “formatting/presentation” problem only and had no impact on the study findings. The originally published article stated correctly that all 55 articles published between issue 3/2009 and 2/2010 were included, but the cited references erroneously contained 12 additional references from issue 2/2009, which were not part of the analysis, for the reasons described in the article (sparse tweetations pre-issue 3/2009). In the corrected version we have not only removed these extra 12 references (31-42), but we also took the opportunity to move all other references of included articles (43-97) into a new Multimedia Appendix 2, no longer citing them in the “References” section. We now refer to them in the paper by article ID (last 4-digits of the DOI), where we previously used in-text citations (Table 2 and Discussion). While there was nothing wrong with the way the articles were cited previously, and while we think that citing the JMIR articles whose impact we discuss in the paper is proper and necessary, we want to avoid any potential impression that this editorial artificially skews JMIR's future impact factor. One way to avoid this is to move the references to a separate file. The original decision to cite them as references was made for the sake of convenience for our readers, to prevent them from having to look up the references in a separate file or by DOI. JMIR has no space limitations and generally prefers to cite references in the article rather than in an Appendix; for readers downloading a PDF file it is more convenient to have all references in a single file rather than having to download a separate Appendix. The decision to now move these references into a Multimedia Appendix was made after a reader and publishing colleague pointed out that citing these articles may increase JMIR’s impact factor. Although none of the two peer-reviewers, both experts in scientometrics, were originally concerned about citing the included articles as references, and even though any potential additional impact factor points after the decimal point caused by the original editorial would probably have been neglible (after all, these articles are already highly cited: altogether, 638 times, according to Google Scholar), and even though Thomson Reuters also publishes a journal impact factor that excludes journal self-citations, we wish to avoid any potential debate or uncertainty on what proportion of future JMIR impact factors were caused by this editorial, and have therefore decided to pre-emptively move these references into a separate file (Multimedia Appendix 2). The article correction was made on January 4, 2012, before submission to PubMed Central, Swets and other content aggregators and databases, and before indexing by Thomson Reuters. Having to remove references from a manuscript to preserve the validity of a journal-level impact metric is somewhat troubling, but if anything, then this perhaps illustrates the limitations and tyranny

of the impact factor, and why we should consider additional metrics.

Edited by G Eysenbach; submitted 04.01.12; this is a non–peer-reviewed article;accepted 04.01.12; published 04.01.12.

Please cite as:
Eysenbach G
Correction: Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact
J Med Internet Res 2012;14(1):e7
URL: http://www.jmir.org/2012/1/e7/
doi:10.2196/jmir.2041
PMID:22222952

Corrigenda and Addenda

Correction: Improving the Quality of Web Surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES)

Gunther Eysenbach[1,2,3], MD, MPH, FACMI

[1]University Health Network, Centre for Global eHealth Innovation & Techna Institute, Toronto, ON, Canada
[2]Institute for Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada
[3]JMIR Publications Inc., Toronto, ON, Canada

Corresponding Author:
Gunther Eysenbach, MD, MPH, FACMI
University Health Network
Centre for Global eHealth Innovation & Techna Institute
190 Elizabeth St
Toronto, ON, M4L3Y7
Canada
Phone: 1 416 7866970
Fax: 1 416 340 3595
Email: geysenba@uhnres.utoronto.ca

Related Article:

Correction of: http://www.jmir.org/2004/3/e34

(J Med Internet Res 2012;14(1):e8) doi:10.2196/jmir.2042

An error in the CHERRIES statement has been corrected (J Med Internet Res 2004;6[3]:e34). In the original paper, in table 1, under the recommendations on how response rates (view rate, participation rate, and completion rate) should be calculated, denominators and numerators were flipped. The view rate should be the ratio of unique survey visitors divided by unique site visitors. The participation rate should be the ratio of those who agreed to participate divided by unique first survey page visitors. The completion rate is the ratio of the number of people who finished the survey divided by those who agreed to participate. The corrections have been made in the table in both columns. A corrected version has been submitted to PubMed Central, but incorrect versions may exist on other sites.

Edited by G Eysenbach; submitted 04.01.12; this is a non–peer-reviewed article;accepted 04.01.12; published 04.01.12.

Please cite as:
Eysenbach G
Correction: Improving the Quality of Web Surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES)
J Med Internet Res 2012;14(1):e8
URL: http://www.jmir.org/2012/1/e8/
doi:10.2196/jmir.2042
PMID:22223031

Corrigenda and Addenda

Metadata Correction: Online Social Networks and Smoking Cessation: A Scientific Research Agenda

Nathan K Cobb[1,2,3,4], MD; Amanda L Graham[1,3], PhD; M. Justin Byron[4], MHS; Raymond S Niaura[1,3,4], PhD; David B Abrams[1,3,4], PhD; Workshop Participants[5]

[1]The Schroeder Institute for Tobacco Research and Policy Studies, American Legacy Foundation, Washington, DC, United States
[2]Division of Pulmonary & Critical Care, Department of Medicine, Georgetown University Medical Center, Washington, DC, United States
[3]Department of Oncology, Georgetown University Medical Center / Lombardi Comprehensive Cancer Center, Washington, DC, United States
[4]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States
[5]see acknowledgements of original article

Corresponding Author:
Nathan K Cobb, MD
The Schroeder Institute for Tobacco Research and Policy Studies
American Legacy Foundation
1724 Massachusetts Ave NW
Washington, DC, 20036
United States
Phone: 1 202 454 5745
Fax: 1 202 454 5785
Email: ncobb@legacyforhealth.org

Related Article:

Correction of: http://www.jmir.org/2011/4/e119

(J Med Internet Res 2012;14(1):e12) doi:10.2196/jmir.2049

Authorship Correction

The authors of "Online Social Networks and Smoking Cessation: A Scientific Research Agenda" (J Med Internet Res 2011;13[4]:e119) inadvertently omitted Raymond S. Niaura from the list of authors during the submission process. The author Niaura should have been added after M. Justin Byron in the originally published manuscript. This error has been corrected in the online version of the paper on the JMIR website on January 11, 2012, together with publishing this correction notice. This was done before submission to Pubmed Central and other full-text repositories, but after submission to PubMed/Medline. Correction of the Pubmed/Medline record will be requested and will be done by the National Library of Medicine (NLM). The Crossref/DOI record has been corrected by the publisher.

Edited by G Eysenbach; submitted 11.01.12; this is a non–peer-reviewed article;accepted 11.01.12; published 11.01.12.

Please cite as:
Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants
Metadata Correction: Online Social Networks and Smoking Cessation: A Scientific Research Agenda
J Med Internet Res 2012;14(1):e12
URL: http://www.jmir.org/2012/1/e12/
doi:10.2196/jmir.2049
PMID:22236454

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org